

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 06:19:20 ; Search time 6510 Seconds
(without alignments)
7700.013 Million cell updates/sec

Title: US-09-597-920B-1

Perfect score: 1060

Sequence: 1 gactctgcctcctgaggggccc.....aaaaaaaaaaaaaaaaaaaaa 1060

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba:*
- 2: gb_hcg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_rtg:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1060	100.0	1060	9 AF036905	AF036905 Homo sapi
2	1027.4	96.9	1616	9 AF036905	AJ223280 Homo sapi
3	1013.4	95.6	1472	9 BC011563	BC011563 Homo sapi
4	949.2	89.5	1244	6 BD107841	BD107841 36 human
5	932	87.9	1460	9 AF036906	AF036906 Homo sapi
6	479.6	45.2	2443	9 AK093815	AK093815 Homo sapi
7	416.8	39.3	528	6 AR380020	AR380020 Sequence
8	413.8	39.0	1260	10 BC052340	BC052340 Mus muscu
9	400.4	37.8	1260	10 AF036907	AF036907 Mus muscu
10	396	37.4	1188	10 RND001184	AJ001184 Rattus no
11	395.4	37.3	1232	10 BC013337	BC013337 Mus muscu
12	265	25.0	131839	9 AC112166	AC112166 Homo sapi
13	265	25.0	163424	9 AC109460	AC109460 Homo sapi
14	265	25.0	187788	2 AC138925	AC138925 Homo sapi
15	265	25.0	219646	2 AC139247	AC139247 Homo sapi
16	205.2	19.4	400	6 CQ664157	CQ664157 Sequence
17	200.6	201	11 BV204291	BV204291 sqm21471	
18	100.6	9.5	182032	2 AC150435	AC150435 Atelesrix
19	67.4	6.4	125020	9 AF429315	AF429315 Homo sapi

20	59.4	5.6	125020	9 AF429315	AF429315 Homo sapi
21	56.2	5.3	140031	2 AC121686	AC121686 Rattus no
22	56.2	5.3	231198	2 AC126892	AC126892 Rattus no
23	54.6	5.2	6307	6 AX818044	AX818044 Sequence
24	54.6	5.2	6307	6 AX826982	AX826982 Sequence
25	54.6	5.2	6307	10 NMU438435	NMU438435 Mus muscu
26	54.6	5.2	217912	10 AC125322	AC125322 Mus muscu
27	53.8	5.1	101451	2 AC138105	AC138105 Mus muscu
28	53.2	5.0	2000	6 AX555393	AX555393 Sequence
29	53	5.0	189190	2 AC129536	AC129536 Mus muscu
30	52.6	5.0	2030	9 BC041637	BC041637 Homo sapi
31	51.4	4.8	7218	6 I66494	I66494 Sequence 14
32	51.2	4.8	101451	2 AC138105	AC138105 Mus muscu
33	51	4.8	218780	2 AC115552	AC115552 Rattus no
34	48.6	4.6	179886	2 AC140072	AC140072 Mus muscu
35	48.6	4.6	219568	10 AL844840	AL844840 Mouse DNA
36	48.4	4.6	162738	9 AC083812	AC083812 Homo sapi
37	48.4	4.6	241432	10 AL589661	AL589661 Mouse DNA
38	48	4.5	2285	9 AK000291	AK000291 Homo sapi
39	48	4.5	231462	2 AC121752	AC121752 Rattus no
40	47.6	4.5	194018	10 AC124686	AC124686 Mus muscu
41	47.4	4.5	1610	8 AK109707	AK109707 Oryza sat
42	47.4	4.5	3196	6 BD232172	BD232172 Novel sec
43	47.4	4.5	123954	8 AC123518	AC123518 Oryza sat
44	47.4	4.5	139943	2 AC119190	AC119190 Mus muscu
45	47.4	4.5	162311	8 AC145321	AC145321 Oryza sat

ALIGNMENTS

RESULT 1	AF036905	1060 bp	mRNA	linear	PRI 06-MAR-1998
LOCUS	AF036905				
DEFINITION	Homo sapiens linker for activation of T cells (LAT) mRNA, complete cds.				
ACCESSION	AF036905				
VERSION	AF036905.1	GI:2828023			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE	Homo sapiens				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
TITLE	1 (bases 1 to 1060)				
REFERENCE	LAT: the ZAP-70 tyrosine kinase substrate that links T cell receptor to cellular activation				
JOURNAL	Cell 92 (1), 83-92 (1998)				
MEDLINE	98149346				
PUBMED	9489702				
REFERENCE	2 (bases 1 to 1060)				
AUTHORS	Zhang, W., Sloan-Lancaster, J., Kitchen, J., Trible, R.P. and Samelson, L.E.				
TITLE	Direct Submission				
JOURNAL	Submitted (05-DEC-1997) Cell Biology and Metabolism Branch, National Institute of Child Health and Development, National Institute of Health, 9000 Rockville Pike, Bethesda, MD 20892, USA				
COMMENT	LAT is a highly tyrosine phosphorylated protein, previously described as p36-38, and it associates with many signaling molecules, such as Grb2, PLC-gamma1, PI-3 kinase, cbl, Vav, and SLP-76, either directly or indirectly upon T cell activation. It is a potential type III transmembrane protein.				
FEATURES	location/Qualifiers				
source	1..1060				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
gene	/cell_line="Jurkat T cells"				
	1..1060				
	/gene="LAT"				
	/note="linker for activation of T cells"				
CDS	58..759				

```

/gene="LAT"
/nc="tyrosine kinase substrate"
/codon_start=1
/product="LAT"
/proteln_id="AAC39636.1"
/db_xref="GI:2828024"
/translation="MEBALIVPCVLGILLPLIAMLALCVHCHRLPGSDYSTSDSL
YPRGIQKRPHTVAWPWPAPVPTVSYPPLSQDPLPIPRSPQLGSHRTPSSRSD
GANSVASYENEPACEDADEDDYHNGYLVLPDSTPATSTPAASAPALSTPGIRD
SAFMSIEDYVNVVPESGESABASLDGSRVNVVSQELHPAAKTEPALASQAEV
EEBAPYENLQELN"

```

ORIGIN

Query Match 100.0%; Score 1060; DB 9; Length 1060;
 Best Local Similarity 100.0%; Pred. No. 7.9e-257;
 Matches 1060; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GACTCTGCCCTTTAGAGGGGCTTAGGGGTGACAGCCCTGCTCCAGCTCCCTGCAATG 60
DB 1 GACTCTGCCCTTTAGAGGGGCTTAGGGGTGACAGCCCTGCTCCAGCTCCCTGCAATG 60
QY 61 GAGAGGGCCATCTGGTCCCTGGGTGCTGGGGGCTCCGTGCTGCCCCCATCTGGCCATG 120
DB 61 GAGAGGGCCATCTGGTCCCTGGGTGCTGGGGGCTCCGTGCTGCCCCCATCTGGCCATG 120
QY 121 TTGATGGCCTGTGTGTGTCACCTGACACAGCTGACAGGCTCTTACGACAGACATCTCTCA 180
DB 121 TTGATGGCCTGTGTGTGTCACCTGACACAGCTGACAGGCTCTTACGACAGACATCTCTCA 180
QY 181 GATAGTTTGTATCCAAAGGGGATCCAGTTCAAAGGGCTCACAGGTTGGCCCTGAGGCA 240
DB 181 GATAGTTTGTATCCAAAGGGGATCCAGTTCAAAGGGCTCACAGGTTGGCCCTGAGGCA 240
QY 241 CCTGCTACCCACCTGTACCTCTTACCCACCTGACAGCCAGCCAGACCTGCTCCCATC 300
DB 241 CCTGCTACCCACCTGTACCTCTTACCCACCTGACAGCCAGCCAGACCTGCTCCCATC 300
QY 301 CCAAGATCCCGCAGACCCCTTGGGGGCTCCAGCCGAGCCGATCTTCCCGGCGGATTTCT 360
DB 301 CCAAGATCCCGCAGACCCCTTGGGGGCTCCAGCCGAGCCGATCTTCCCGGCGGATTTCT 360
QY 361 GATGCTGCAACAGTGTGGGAGCTACGAAACGAGAAACGAGCTGTGAGAGATGCAGAT 420
DB 361 GATGCTGCAACAGTGTGGGAGCTACGAAACGAGAAACGAGCTGTGAGAGATGCAGAT 420
QY 421 GAGATGAGAGAGCTATCACAAACCCAGGCTACCGTGTGTGCTTCTTGACAGCACCCCG 480
DB 421 GAGATGAGAGAGCTATCACAAACCCAGGCTACCGTGTGTGCTTCTTGACAGCACCCCG 480
QY 481 GCCACTAGCACTGTGCCCCCATCACTGCTGACACAGCCCTGGCATCCGAGACAGT 540
DB 481 GCCACTAGCACTGTGCCCCCATCACTGCTGACACAGCCCTGGCATCCGAGACAGT 540
QY 541 GCTCTTCCATGAGAGTCCATTGATGATTAAGTAAAGTTCGAGAGAGCGGAGAGCCCA 600
DB 541 GCTCTTCCATGAGAGTCCATTGATGATTAAGTAAAGTTCGAGAGAGCGGAGAGCCCA 600
QY 601 GAAGGCTCTGAGTGGAGCGGAGGATGAGATGTCCTCCAGAACTGATCTCTGGA 660
DB 601 GAAGGCTCTGAGTGGAGCGGAGGATGAGATGTCCTCCAGAACTGATCTCTGGA 660
QY 661 GCGGCTAAGTGAAGCTGCGCCCTGAGTTCCAGAGAGGAGAGAGAGTGAAGAG 720
DB 661 GCGGCTAAGTGAAGCTGCGCCCTGAGTTCCAGAGAGGAGAGAGTGAAGAG 720
QY 721 GGGGCTCCAGATTAAGAGATCTGAGAGAGCTGAAGCTGAGGCTGTGAGAGCGGAGTCT 780
DB 721 GGGGCTCCAGATTAAGAGATCTGAGAGAGCTGAAGCTGAGGCTGTGAGAGCGGAGTCT 780
QY 781 GTCCGGAACAGAGCTTGTGCGGAGCGGCTGAGCTGGGCACTGTGAAATGTGCTGTGGGCT 840
DB 781 GTCCGGAACAGAGCTTGTGCGGAGCGGCTGAGCTGGGCACTGTGAAATGTGCTGTGGGCT 840

```

```

QY 841 CTTACATGCGGCTCTGCTGCTGCTCCAGCCTTGACAAACAGCTGAGAAATCCCCCGTAA 900
DB 841 CTTACATGCGGCTCTGCTGCTGCTCCAGCCTTGACAAACAGCTGAGAAATCCCCCGTAA 900
QY 901 CTTATTTATCCTTTGGGGGTTGCGGCTGTGTCCCGGAGGCTGTGCACTTTGAGAGCAG 960
DB 901 CTTATTTATCCTTTGGGGGTTGCGGCTGTGTCCCGGAGGCTGTGCACTTTGAGAGCAG 960
QY 961 CTTGAGAAATGACCTGCGCCCTGAGCCAGCCCTTACTGTGTGTAATGAATAAAGGCTGCGT 1020
DB 961 CTTGAGAAATGACCTGCGCCCTGAGCCAGCCCTTACTGTGTGTAATGAATAAAGGCTGCGT 1020
QY 1021 GTGCTGTGGAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1060
DB 1021 GTGCTGTGGAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1060

```

RESULT 2
 HSAJ3280
 LOCUS
 DEFINITION Homo sapiens mRNA for 36 kDa phosphotyrosine protein.
 ACCESSION AJ223280
 VERSION AJ223280.1 GI:2780222
 KEYWORDS phosphotyrosine protein.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
 AUTHORS Weber, J.R., Orstavik, S., Torgersen, K.M., Danbolt, N.C., Berg, S.F., Ryan, J.C., Tasken, K., Imboden, J.B. and Vaae, J.T.
 TITLE Molecular cloning of the cDNA encoding pp36, a tyrosine-phosphorylated adaptor protein selectively expressed by T cells and natural killer cells
 JOURNAL J. Exp. Med. 187 (7), 1157-1161 (1998)
 MEDLINE 98197173
 PUBMED 9529333

REFERENCE 2 (bases 1 to 1616)
 AUTHORS Orstavik, S.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Orstavik S., Institute of Medical Biochemistry, University of Oslo, BP 1112, Blindern, N-0317 Oslo, NORWAY

FEATURES

source location/Qualifiers

gene
 1..1616
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="thymus"

CDS
 1..1616
 /gene="pp36"
 323..1024
 /gene="pp36"

ORIGIN
 Query Match 96.9%; Score 1027.4; DB 9; Length 1616;
 Best Local Similarity 99.9%; Pred. No. 1.4e-248;
 Matches 1026; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 GACTCTGCCCTTTAGAGGGGCTTAGGGGTGACAGCCCTGCTCCAGCTCCCTGCAATG 60
DB 266 GACTCTGCCCTTTAGAGGGGCTTAGGGGTGACAGCCCTGCTCCAGCTCCCTGCAATG 325

```



```

gene
1. .1472
   /gene="LAT"
   /db_xref="LocustID:27040"
   /db_xref="MIM:602354"
   145. .846
   /gene="LAT"
   /codon_start=1
   /product="LAT protein"
   /protein_id="AAH1563.1"
   /db_xref="GI:15079451"
   /db_xref="LocustID:27040"
   /db_xref="MIM:602354"
   /translation="MEBALIVPCVGLGLPLPIALMALCVHCHRLPGSYDSTSSDIL
YRGIOFKPHTVAPMPAPVAPVPTSGYPTLSODLLIPRSPQIGSGHRTSPSSRSD
GANSVAAYNEERACEDADEDEDHNGCYIVLPDSTPATSTAPASAPALSTGCIID
SAFMSHSIDIVYVPESSGSSAPASLIDSGSREIVYVNSQELHPGAKTERPALSSQAEERY
EEEGAPDENYLOELN"

```

ORIGIN

Query Match	95.6%	Score 1013.4;	DB 9;	Length 1472;
Best Local Similarity	99.9%	Pred. No. 4.8e-245;		
Matches 1014; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	15	GGGGCTTGGGGGTGACGACGACCTCTGCTCCAGCTCCCTGCAGATGAGAGGCCATCT	74
Db	102	GGGGCTTGGGGGTGACGACGACCTCTGCTCCAGCTCCCTGCAGATGAGAGGCCATCT	161
QY	75	GGTCCCTGCGTGTGGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	134
Db	162	GGTCCCTGCGTGTGGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	221
QY	135	TGTGCACTGCGACAGCTGCGAGGGTCTTAAGACAGCAATCTCCAGATGTTGTATTC	194
Db	222	TGTGCACTGCGACAGCTGCGAGGGTCTTAAGACAGCAATCTCCAGATGTTGTATTC	281
QY	195	AAGGGGCACTCCAGTTCAAACGGCTCAGACGGTTGCCCCCTTGCGACCTGCTCAACCA	254
Db	282	AAGGGGCACTCCAGTTCAAACGGCTCAGACGGTTGCCCCCTTGCGACCTGCTCAACCA	341
QY	255	TGTCACTCTCTAACCAACCCCTGAGCTGAGCGACAGACCTGCTCCCATCCAAAGTCCGCA	314
Db	342	TGTCACTCTCTAACCAACCCCTGAGCGACAGACCTGCTCCCATCCAAAGTCCGCA	401
QY	315	GCCCCCTTGGGGGGCTCCCAACGGAGCCACTTCCGGGGGGGATTTGTAGTGGGCCAAG	374
Db	402	GCCCCCTTGGGGGGCTCCCAACGGAGCCACTTCCGGGGGGGATTTGTAGTGGGCCAAG	461
QY	375	TGTGGCGAGCTACGAGAAAGAGAAACAGACCTGTGAGATGACATGAGATGAGAGAGA	434
Db	462	TGTGGCGAGCTACGAGAAAGAGAAACAGACCTGTGAGATGACATGAGATGAGAGAGA	521
QY	435	CTATTCACAAACCAAGGCTACTGTGTGTGTCTTCTCAACAGCACCCCGGCCACTAGCATGC	494
Db	522	CTATTCACAAACCAAGGCTACTGTGTGTGTCTTCTCAACAGCACCCCGGCCACTAGCATGC	581
QY	495	TGCCCATCAGCTCCCTGCACTCAGACCCCTGGGCATCCGAGACATGTGCTTCCATGGA	554
Db	582	TGCCCATCAGCTCCCTGCACTCAGACCCCTGGGCATCCGAGACATGTGCTTCCATGGA	641
QY	555	GTCACATTGATGATTAAGTGAACGTTCCGAGAGACGGGAGAGACCGAGAACGCTCTTGA	614
Db	642	GTCACATTGATGATTAAGTGAACGTTCCGAGAGACGGGAGAGACCGAGAACGCTCTTGA	701
QY	615	TGGCAGCCGGGAGTATGTGAATGTGTCCCAAGAACTGCATCTGTGAAGGGGCTTAACTGA	674
Db	702	TGGCAGCCGGGAGTATGTGAATGTGTCCCAAGAACTGCATCTGTGAAGGGGCTTAACTGA	761
QY	675	GCTTCGCCCCCTGAGTTCCACAGAGGACAGAGAACTGAGAGAAAGGGGGCTCCAGATT	734
Db	762	GCTTCGCCCCCTGAGTTCCACAGAGGACAGAGAACTGAGAGAAAGGGGGCTCCAGATT	821
QY	735	CGAGATCTGACAGAGCTGAACCTGAGGGCTGTGAGAGCCGAGTCTGTCTTGAACCAAG	794

Db	822	CGAATACTCGAGGAGCTGAACTGAGGGGCTGTGTGAGGGCCGAGCTGTCTCTGGAAACGAG	881
Oy	795	CTTGCCCTTGGGACGGCTGAGCTGTGGGCACCTGGAAAGTGGCTCTCTGGGGTCTCAACAGGGGCTC	854
Db	882	CTTGCCCTTGGGACGGCTGAGCTGTGGGCACCTGGAAAGTGGCTCTCTGGGGTCTCAACAGGGGCTC	941
Oy	855	CTGCGCTTGTGCTCCAGCGCTGACCAACAGCGCTGAGAAATCCCCCGTAACTTTATTACACTTT	914
Db	942	CTGCGCTTGTGCTCCAGCGCTGACCAACAGCGCTGAGAAATCCCCCGTAACTTTATTACACTTT	1001
Oy	915	GGGGTTTGGGCTGTGTCTCCCGGAAAGCTCTTGACACTTCTGACGCAAGCTTGAGAAATGACT	974
Db	1002	GGGGTTTGGGCTGTGTCTCCCGGAAAGCTCTTGACACTTCTGACGCAAGCTTGAGAAATGACT	1061
Oy	975	GGCCCTGGCCCCAGGCCACTGCTGTGTAATGAAATAAAGGCTGGGCTGTGCTGTG	1029
Db	1062	GGCCCTGGCCCCAGGCCACTGCTGTGTAATGAAATAAAGGCTGGGCTGTGCTGTG	1116

RESULT 4				
BD107841	BD107841	1224 bp	DNA	linear
LOCUS	36 human secreted proteins.			PAT 18-SEP-2002
DEFINITION				
ACCESSION	BD107841			
VERSION	BD107841.1 GI:23202659			
KEYWORDS	JF 2002500035-A/I2.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			

REFERENCE 1 (bases 1 to 1224)
AUTHORS Ruben,S.M., Soppet,D.R., Ebner,R., Lafleur,D.W., Ni,J.,
Brewer,L.A., Olsen,H.S., Duan,R.D. and Rosen,C.A.
TITLE 36 human secreted proteins
JOURNAL Patent: JP 2002500035-A 12 08-JAN-2002;
COMMENT HUMAN GENOME SCIENCES INC
OS Homo sapiens (human)

```

PF      06-JAN-1998 JP 2000527554
PR      07-JAN-1998 US 60/070657,07-JAN-1998 US 60/070658 PR
07-JAN-1998 US 60/070652,07-JAN-1998 US 60/070704 PI STEVEN
M RUBEN ,DANIEL R SOPPET,REINHARD EBNER,DAVID W LAFFLEUR, PI JIAN
NI,
PI      LAURIE A BREWER,HENRIK S OLSEN,ROSANNE D DUAN,CRAIG A ROSEN PC
C12N15/09,A61K31/711,A61K38/00,A61K39/395,A61K39/395,A61K48/00,
PC      A6195/00,
PC      A61P7/00,A61P11/06,A61P19/02,A61P29/00,A61P31/18,A61P35/00, PC
A61P35/02,
PC      A61P37/00,C07K14/435,C07K16/18,C12N1/15,C12N1/21,C12N5/10, PC
C12N15/00,
PC      A61K37/02,C12N5/00
CC      n equale a,t,s, or c
FH      Key location/Qualifiers
FT      source 1..1224
FT      /Organism='Homo sapiens (human)'.
FEATURES
source 1..1224

```

ORIGIN

Query Match	89.5%	Score 949.2;	DB 6;	Length 124;
Best Local Similarity	98.9%	Pred. No. 8e-229;		
Matches 1039; Conservative	0;	Mismatches	4;	Indels 8; Gaps 8;

Oy I GACTCTGCCCTTGAGGGGGCTGACAGCCAGCTCTCGAGCTCCCCGCAATG 60
 Db 181 GACTCTGCCCTTGAGGGGGCTGACAGCCAGCTCTCGAGCTCCCCGCAATG 240
 Oy 61 GAGGAGGCATCTCTGATCCCCCTGCTGCTGGGGCTCTGCTGCTGCCATCTGGCATG 120

Db 241 GAGGAGCCATCTGTCCTCCCTGCGTGTGGGGCTCTCTGCTGCGCATCTCTGGCCATG 300
 Qy 121 TTGATGGCACTGTGTGTGCACTGCCACAGAGCTCTCTAGAGACGACATCTCTCA 180
 Db 301 TTGATGGCACTGTGTGTGCACTGCCACAGAGCTCTCTAGAGACGACATCTCTCA 360
 Qy 181 GATAGTTGTATCC-AAGGGGCACTCAGTTCAACAGGCTCTCAACAGGTTGCCCTGAGCC 239
 Db 361 GATAGTTGTATCCAAAGGGGCACTCAGTTCAACAGGCTCTCAACAGGTTGCCCTGAGCC 420
 Qy 240 ACTGCGCTACCACTCTGTCACTCTCACTCCACCTCTGAGCCAGCCAGACTCTCTCCCAT 299
 Db 421 ACTGCGCTACCACTCTGTCACTCTCACTCCACCTCTGAGCCAGCCAGACTCTCTCCCAT 480
 Qy 300 CCCAAGATCCCGGAGCCCTTGGGGGCTCCCAACCGGAGCCATCTTCCCGGCGGATTC 359
 Db 481 CCCAAGATCCCGGAGCCCTTGGGGGCTCCCAACCGGAGCCATCTTCCCGGCGGATTC 540
 Qy 360 TGATGGTGCACAGTGTGCGAGCTACGAGAACGAGAACAGCCCTGTGAGAGTCAGA 419
 Db 541 TGATGGTGCACAGTGTGCGAGCTACGAGAACGAGAACAGCCCTGTGAGAGTCAGA 599
 Qy 420 TGAGATGAGAGCACTATCACAAACCAGGCTAAGCTGTGTCTTCTGACAGCACTCC 479
 Db 600 TGA-GATGAGAGCACTATCACAAACCAGGCTAAGCTGTGTGTCTTCTGACAGCACTCC 658
 Qy 480 GGGCACTAGCACTGTGCTGCTCCATCAGCTCTGTGCACTGACAGACCTCTGGCATCCGAGACG 539
 Db 659 GGGCACTA-CAGTGTGCTGCTCCATCAGCTCTGTGCACTGACAGACCTCTGGCATCCGAGACG 717
 Qy 540 TGCTTCTTCATGAGTGCATGATGATTAAGTGAAGCTTCCGAGAGCGGGAGAGAGCC 599
 Db 718 TGCTTCTTCATGAGTGCATGATGATTAAGTGAAGCTTCCGAGAGCGGGAGAGAGCC 777
 Qy 600 AGAAGCGTCTGTGATGAGGAGCGGAGATGATGTGTCCAGGAATCTGCATCTCTG 659
 Db 778 AGAAGCGTCTGTGATGAGGAGCGGAGATGATGTGTCCAGGAATCTGCATCTCTG 837
 Qy 660 AGCGCTTAAGATGAGTGCCTGCGCTGAGTTCCAGAGAGGAGAGAGTGTGAGAGAGA 719
 Db 838 AGCGCTTAAGATGAGTGCCTGCGCTGAGTTCCAGAGAGGAGAGAGTGTGAGAGAGA 897
 Qy 720 GGGGGCTCCAGATTCAGAGATCTGAGAGGAGCTGAGAGGCTGTGAGAGCGGAGTC 779
 Db 898 GGGGGCTCCAGATTCAGAGATCTGAGAGGAGCTGAGAGGCTGTGAGAGCGGAGTC 957
 Qy 780 TGTCTGTGAAACAGGCTGTGCTGTGAGAGCGGCTGAGCTGAGAGTGTGCTGTGAGG 839
 Db 958 TGTCTGTGAAACAGGCTGTGCTGTGAGAGCGGCTGAGCTGAGAGTGTGCTGTGAGG 1017
 Qy 840 TCTTCAATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 899
 Db 1018 TCTTCAATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1077
 Qy 900 ACTTATTTATCACTTTGGGGTTGGGGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 959
 Db 1078 ACTTATTTATCACTTTGGGGTTGGGGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1135
 Qy 960 GCTTGAAGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1019
 Db 1136 GCTTGAAGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1193
 Qy 1020 TGTGTCTGTGAAAAAATTT 1050
 Db 1194 TGTGTCTGTGTAAAAAATTT 1224

RESULT 5
 AF036906 1460 bp mRNA linear PRI 06-MAR-1998
 LOCUS AF036906
 DEFINITION Homo sapiens linker for activation of T cells (LAT) mRNA,
 alternatively spliced form, complete cds.
 ACCESSION AF036906

VERSION AF036906.1 GI:2828025
 KEYWORDS Homo sapiens (human)
 SOURCE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 ORGANISM Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1460)
 AUTHORS Zhang, W., Sloan-Lancaster, J., Kitchen, J., Tribble, R.P. and
 Samelson, L.E.
 TITLE LAT: the ZAP-70 tyrosine kinase substrate that links T cell
 JOURNAL receptor to cellular activation
 MEDLINE Cell 92 (1), 83-92 (1998)
 PUBMED 9489702
 REFERENCE 2 (bases 1 to 1460)
 AUTHORS Zhang, W., Sloan-Lancaster, J., Kitchen, J., Tribble, R.P. and
 Samelson, L.E.
 TITLE Direct Submision
 JOURNAL Submitted (05-DEC-1997) Cell Biology and Metabolism Branch,
 National Institute of Child Health and Development, National
 Institute of Health, 9000 Rockville Pike, Bethesda, MD 20892, USA
 COMMENT LAT is a highly tyrosine phosphorylated protein, previously
 described as p36-38, and it associates with many signaling
 molecules, such as Grb2, PLC-gamma1, PI-3 kinase, cbl, Vav, and
 SLP-76, either directly or indirectly upon T cell activation. It is
 a potential type III transmembrane protein.
 FEATURES
 source location/Qualifiers
 1..1460
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_line="Jurkat T cells"
 1..1460
 /gene="LAT"
 /note="Linker for activation of T cell"
 79..867
 /gene="LAT"
 /note="tyrosine kinase substrate; This is alternatively
 spliced form of LAT"
 /codon_start=1
 /product="LAT"
 /protein_id="AAC39637.1"
 /db_xref="GI:2828026"
 /translation="MEEALVPCVGLILPLIAMLALCVHCHRLPSVYSTSDSL
 YPRGIQKRPHTVAPWPPAYPPVTSYPLSPDLIPRSPQPLGSGRTSSRRSD
 GANSVASYLENGASGIRGAQAGWGWGWSWTRLPVSLPPPACEDADEDDYHNPG
 YLVVLPSSTPATSTPAESPALSTPGIDSAFMSISIDYVNVFSGRSABASLDGSR
 EYVNVSGRLHGAATKEBPALSSQEAEEVEEAGAPDYENLQELN"
 ORIGIN
 Query Match 87.9%; Score 932; DB 9; Length 1460;
 Best Local Similarity 92.2%; Pred. No. 1.8e-224;
 Matches 1029; Conservative 0; Mismatches 0; Indels 87; Gaps 1;

Qy 1 GACTTGTGCTTGAAGGGGCTTAAGGGTGCACGAGCTGTCCGAGCTCCCTCGAGATG 60
 Db 22 GACTTGTGCTTGAAGGGGCTTAAGGGTGCACGAGCTGTCCGAGCTCCCTCGAGATG 81
 Qy 61 GAGGAGGCACTCTGGTCTCCCTGAGTGTGGGGCTCTGCTGCTGCTGCTGCTGCTGCTG 120
 Db 82 GAGGAGGCACTCTGGTCTCCCTGAGTGTGGGGCTCTGCTGCTGCTGCTGCTGCTGCTG 141
 Qy 121 TTGATGGCACTGTGTGTGCACTGCCACAGACTGCCAGGCTCTTACGACAGCATCTCTCA 180
 Db 142 TTGATGGCACTGTGTGTGCACTGCCACAGACTGCCAGGCTCTTACGACAGCATCTCTCA 201
 Qy 181 GATAGTTGTATCCAAAGGGGCACTCAGTTCAAGGCGCTCAAGGCTTGGCCCGCGGCA 240
 Db 202 GATAGTTGTATCCAAAGGGGCACTCAGTTCAAGGCGCTCAAGGCTTGGCCCGCGGCA 261
 Qy 241 CTTGCTTACCACTGTGCACTCTTACCACTCTGAGAGGAGAGAGAGAGAGAGAGAGAGAG 300
 Db 262 CTTGCTTACCACTGTGCACTCTTACCACTCTGAGAGGAGAGAGAGAGAGAGAGAGAGAG 321

QY	301	CCAAGATCCCCGACAGCCCTTGGGGGCTCCCAACGGACCATCTTCCCGGGGATCT	360
Db	322	CCAAATCCCCCGAGCCCTTGGGGGCTCCCAACGGACCATCTTCCCGGGGATCT	381
QY	361	GATGGTGCACACAGTGTGGCGAGCTACGAGAACGAG-----	396
Db	382	GATGGTGCACACAGTGTGGCGAGCTACGAGAACGAGGTTGGCTGTGGATCCGAGTGCC	441
QY	397	-----	396
Db	442	CAGGCTGGGTGGGAGTCTGGGGTTCGTCCTGGACTAGGCTGACCCCTGTGTCTTACC	501
QY	397	---GAAACAGGCTGTGAGAGATGCATGATAGATGAGACGACTATCACAAACCGAGCTAC	453
Db	502	CCAGAACAGGCTGTGAGAGATGCATGATAGATGAGACGACTATCACAAACCGAGCTAC	561
QY	454	CTGGTGTGTCTTCTTGACAGCACCCCGGCCTAGCAGCTGCTGCCCATCAGCTCTTGCA	513
Db	562	CTGGTGTGTCTTCTTGACAGCACCCCGGCCTAGCAGCTGCTGCCCATCAGCTCTTGCA	621
QY	514	CTCAGCAACCCCTGGGATCCGAGACAGTGCCTTCTCAGTGGATGCATGTAGATGATTAAGTG	573
Db	622	CTCAGCAACCCCTGGGATCCGAGACAGTGCCTTCTCAGTGGATGCATGTAGATGATTAAGTG	681
QY	574	AACGTTCCGAGAGAGCGGGGAGNAGCCAGAAAGCGTCTTGATATGACAGCGGGAGATGTG	633
Db	682	AACGTTCCGAGAGAGCGGGGAGNAGCGCAGAAAGCGTCTTGATATGACAGCGGGAGATGTG	741
QY	634	AATGTGTCCCAAGAACTGCAATCTTGAGAGCGCTTAAGACTGAGCCTGCCCTGAGTTCC	693
Db	742	AATGTGTCCCAAGAACTGCAATCTTGAGAGCGCTTAAGAAGTCCGCCGCTTGAGTTCC	801
QY	694	CAGAGAGCAGAGGAAGTGAAGGAAGGGGGCTCCAGATTACAGAAATCTGCAGAGGCTG	753
Db	802	CAGAGAGCAGAGGAAGTGAAGGAAGGGGGCTCCAGATTACAGAAATCTGCAGAGGCTG	861
QY	754	AACTAGGGGCTGTGAGGCGGAGTCTGTCTTGAAACCAAGGCTTGCTGGGACGGCTGAG	813
Db	862	AACTAGGGGCTGTGAGGCGGAGTCTGTCTTGAAACCAAGGCTTGCTGGGACGGCTGAG	921
QY	814	CTGGGCACTGGAAGTGGCTCTGGGGGTCTCAACATGGGGTCTTGCCCTTGCTCCAGGCTG	873
Db	922	CTGGGCACTGGAAGTGGCTCTGGGGGTCTCAACATGGGGTCTTGCCCTTGCTCCAGGCTG	981
QY	874	ACAACAGGCTGAGAAATCCCCCGTAACCTTATTAACCTTGGGGGTTGGGCTGTGTCCC	933
Db	982	ACAACAGGCTGAGAAATCCCCCGTAACCTTATTAACCTTGGGGGTTGGGCTGTGTCCC	1041
QY	934	CCGAAAGCTGTGCACCTTCTGACGAGGCTTGGAATGACCTTGCCCTGGACCCCAAGCCTAC	993
Db	1042	CCGAAAGCTGTGCACCTTCTGACGAGGCTTGGAATGACCTTGCCCTGGACCCCAAGCCTAC	1101
QY	994	TCTGTGTATAAGAAATAAGGCTGTGGTGTGTCTGTG 1029	
Db	1102	TCTGTGTATAAGAAATAAGGCTGTGGTGTGTCTGTG 1137	

RESULT 6	AK093815	LOCUS	AK093815	2443 bp	mRNA	linear	PRI 30-JAN-2004
DEFINITION	Homo sapiens	CDNA	FLJ36496	fib.	clone	THYM2018819.	
ACCESSION	AK093815						
VERSION	AK093815.1	GI:21752752					
KEYWORDS	oligo capping; fib	(full insert sequence).					
SOURCE	Homo sapiens	(human)					
ORGANISM	Homo sapiens						
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;						
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.						
REFERENCE	1	Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakematsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Odayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,					
AUTHORS							

TITLE
Complete sequencing and characterization of 2l,243 full-length human cDNAs

JOURNAL REFERENCE
Nat. Genet. 36 (1), 40-45 (2004)

AUTHORS
Ishibashi, T., Kanehori, K., Yoshida, M., Watanabe, S., Ishida, S., Ono, Y., Hottel, T., Hiraoaka, S., Murakami, K., Takiguchi, S., Kusano, J., Watanabe, M., Fujimori, K., Tanai, H., Ishida, M., Yamashita, H., Chiba, Y., Sugiyama, T., Irie, R., Otsubi, T., Sato, H., Ota, T., Wakamatsu, A., Ishii, S., Yamamoto, D., Isono, Y., Kawai, Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuno, Y., Nagai, K. and Isogai, T.

TITLE
NEBO human cDNA sequencing project

JOURNAL REFERENCE
Unpublished
3 (bases 1 to 2443)

AUTHORS
Isogai, T. and Yamamoto, J.

COMMENT
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kanusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomic@hri.co.jp, Tel:81-438-52-32975, Fax:81-438-52-3966)
NEBO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- and 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB;
location: HRI and RAB.
Location/Qualifiers

FEATURES
source

ORIGIN

Query Match 45.2%; Score 479.6; DB 9; Length 2443;

Best Local Similarity 99.2%; Pred. No. 4,1e-110;

Matches 482; Conservative 0; Mismatches 6; Gaps 0;

544 TTCTCATGGAGTCATTGATGTAGTACCGAAGCTTCCGAGAGCGGAGAGCAGCAAA 603

Db 1955 TCACAGTGGAGTCCATTGATGATTAAGTGAACGTTCCGAGAGCGGGAGAGCCAGAA 2014
 Qy 604 GGTCTCTGGATGAGAGCCGGAGATATGTAATGTGTCCAGAACTGATCTTGAGAGC 663
 Db 2015 GGTCTCTGGATGAGAGCCGGAGATATGTAATGTGTCCAGAACTGATCTTGAGAGC 2074
 Qy 664 GCTAAGACTGAGAGCCGCGCTGAGATTCCAGAGAGCAGAGAAAGTGAAGAGAGG 723
 Db 2075 GCTAAGACTGAGAGCCGCGCTGAGATTCCAGAGAGCAGAGAAAGTGAAGAGAGG 2134
 Qy 724 GCTCAAGATTACGAGAAATCTGAGAGAGCTGAAGAGGCGCTGTGAGAGCGGAGTCTGTC 783
 Db 2135 GCTCAAGATTACGAGAAATCTGAGAGAGCTGAAGAGGCGCTGTGAGAGCGGAGTCTGTC 2194
 Qy 784 CTGGAACAGAGCTTCCCTGGAGAGCGCTGAGCTGGAGAGTGAAGTGGCTTGGGGTCTCT 843
 Db 2195 CTGGAACAGAGCTTCCCTGGAGAGCGCTGAGCTGGAGAGTGAAGTGGCTTGGGGTCTCT 2254
 Qy 844 CACATGCGCTCTGCGCTTGTCTCCAGAGCTGAGAGAGCTGAGAAATCCCGCGTAATT 903
 Db 2255 CACATGCGCTCTGCGCTTGTCTCCAGAGCTGAGAGAGCTGAGAAATCCCGCGTAATT 2314
 Qy 904 ATTATCACTTTGGGGTGTGGCGCTGTGTCTCCGAAAGCTCTGACCTTGTGAGAGCCT 963
 Db 2315 ATTATCACTTTGGGGTGTGGCGCTGTGTCTCCGAAAGCTCTGACCTTGTGAGAGCCT 2374
 Qy 964 GAGATGACCTGCGCTGCGCGCCAGAGCTTACTGTGTATTAAGAAATGAGGCTTGGCTGTG 1023
 Db 2375 GAGATGACCTGCGCTGCGCGCCAGAGCTTACTGTGTATTAAGAAATGAGGCTTGGCTGTG 2434
 Qy 1024 TCTGTG 1029
 Db 2435 TCTGTG 2440

RESULT 7
 AR380020
 LOCUS AR380020 565 bp DNA linear PAT 18-DEC-2003
 DEFINITION Sequence 565 from patent US 6607879.
 ACCESSION AR380020
 VERSION AR380020.1 GI:40087654
 KEYWORDS
 SOURCE
 ORGANISM
 Unclassified.
 1 (bases 1 to 528)
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source
 1. 528
 /organism="unknown"
 /mol_type="genomic DNA"
 ORIGIN
 Query Match 39.3%; Score 416.8; DB 6; Length 528;
 Best Local Similarity 99.3%; Pired. No. 2.8e-94;
 Matches 429; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Db 1955 TCACAGTGGAGTCCATTGATGATTAAGTGAACGTTCCGAGAGCGGGAGAGCCAGAA 2014
 Qy 604 GGTCTCTGGATGAGAGCCGGAGATATGTAATGTGTCCAGAACTGATCTTGAGAGC 663
 Db 2015 GGTCTCTGGATGAGAGCCGGAGATATGTAATGTGTCCAGAACTGATCTTGAGAGC 2074
 Qy 664 GCTAAGACTGAGAGCCGCGCTGAGATTCCAGAGAGCAGAGAAAGTGAAGAGAGG 723
 Db 2075 GCTAAGACTGAGAGCCGCGCTGAGATTCCAGAGAGCAGAGAAAGTGAAGAGAGG 2134
 Qy 724 GCTCAAGATTACGAGAAATCTGAGAGAGCTGAAGAGGCGCTGTGAGAGCGGAGTCTGTC 783
 Db 2135 GCTCAAGATTACGAGAAATCTGAGAGAGCTGAAGAGGCGCTGTGAGAGCGGAGTCTGTC 2194
 Qy 784 CTGGAACAGAGCTTCCCTGGAGAGCGCTGAGCTGGAGAGTGAAGTGGCTTGGGGTCTCT 843
 Db 2195 CTGGAACAGAGCTTCCCTGGAGAGCGCTGAGCTGGAGAGTGAAGTGGCTTGGGGTCTCT 2254
 Qy 844 CACATGCGCTCTGCGCTTGTCTCCAGAGCTGAGAGAGCTGAGAAATCCCGCGTAATT 903
 Db 2255 CACATGCGCTCTGCGCTTGTCTCCAGAGCTGAGAGAGCTGAGAAATCCCGCGTAATT 2314
 Qy 904 ATTATCACTTTGGGGTGTGGCGCTGTGTCTCCGAAAGCTCTGACCTTGTGAGAGCCT 963
 Db 2315 ATTATCACTTTGGGGTGTGGCGCTGTGTCTCCGAAAGCTCTGACCTTGTGAGAGCCT 2374
 Qy 964 GAGATGACCTGCGCTGCGCGCCAGAGCTTACTGTGTATTAAGAAATGAGGCTTGGCTGTG 1023
 Db 2375 GAGATGACCTGCGCTGCGCGCCAGAGCTTACTGTGTATTAAGAAATGAGGCTTGGCTGTG 2434
 Qy 1024 TCTGTG 1029
 Db 2435 TCTGTG 2440

RESULT 8
 BC052340
 LOCUS BC052340 1260 bp mRNA linear ROD 30-JUN-2004
 DEFINITION Mus musculus linker for activation of T cells, mRNA (cDNA clone MGC:60595 IMAGE:30077071), complete cds.
 ACCESSION BC052340
 VERSION BC052340.1 GI:30354172
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1260)
 AUTHORS
 TITLE
 JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 REMARK
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Drs. Dennis Taub, Dan Longo (NIA, USA), Jonathan Keller (NCI, USA)
 cDNA Library Preparation: Yulan Piao and Minoru Ko (National Institute on Aging, NIH: <http://jgsun.grc.nia.nih.gov/cDNA/>)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
 Galtherburg, Maryland;

AUTHORS Zhang, W., Sloan-Lancaster, J., Kitchen, J., Tribble, R. P. and Samelson, L. E.

TITLE Direct Submission
JOURNAL Submitted (05-DEC-1997) Cell Biology and Metabolism Branch, National Institute of Child Health and Development, National Institute of Health, 9000 Rockville Pike, Bethesda, MD 20892, USA

COMMENT

LAT is a highly tyrosine phosphorylated protein, previously described as p36-38, and it associates with many signaling molecules, such as Grb2, Plc-gamma1, PI-3 kinase, cbl, Vav, and SLP-76, either directly or indirectly upon T cell activation. It is a potential type III transmembrane protein.

FEATURES

source

1..1260

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/tissue_type="thymus"

/dev_stage="adult"

1..1260

/gene="LAT"

/note="linker for activation of T cells"

292..1020

/gene="LAT"

/note="tyrosine kinase substrate"

/codon_start=1

/product="LAT"

/protein_id="AAC40054.1"

/db_xref="GI:2828028"

/translation="MEADALSPVGLGLLPLVLTLLALCYRCRELPVSYDSTESLYPRSLIKPQITVPRTPAVSPLPSPROPDLPIPRSPQGLSHMPSQNSDDANVASVYENQSPACKNVDDEDIDYPNGYLVLPDSSPAAPVVSAPVSNPDIGDAFVSCEDEDYVNPESBESAASLDSREYVNSPEQQPTRALIASVNSQEV EDEGESEGVDSERAPDYENQELN"

ORIGIN

Query Match 37.8%; Score 400.4; DB 10; Length 1260;

Best Local Similarity 67.6%; Pred. No. 4.1e-90;

Matches 698; Conservative 0; Mismatches 291; Indels 43; Gaps 8;

```

QY 26 GTGCAGCCAGCTGCTCCGAGCTCCCTGCAATGAGAGAGCCATCTGTCCTGCG 85
DB 260 GTGCAGAGAGGCGCTGTGCGAGGCTCTGCAATGAGAGAGCCGCTTGAGCCGCGTG 319
QY 86 TGCTGGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 142
DB 320 GCGTGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 379
QY 143 GCCAGAGCTGCGAGGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 202
DB 380 GCGGAGAGTGGCAGAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 439
QY 203 TCCAGTTCAACGGGCTCAACAGGTTGCCCTGAGCAGCTG-----CTTACCCAGCTG 256
DB 440 TCCATATCAAGCAGCTCAATTAACCGTCCCGGAGACAGCTGTTCTTCTTACCTCTAG 499
QY 257 TACCTCTTCAACCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 316
DB 500 TCACTTCTTCCACCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 559
QY 317 CCCTTGGGGGCTCCCAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 376
DB 560 CCCTTGGGGGCTCCCAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 619
QY 377 TGGCAGAGTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 436
DB 620 TGGCAAGTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 679
QY 437 ATCAACAACCC---AGCTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 493
DB 680 AGGACTATCCCAACGGCTACTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 739
QY 494 CTGCCCATCAGCTCTGTGACTCAGACCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 553

```

```

DB 740 TTGTCTCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 799
QY 554 AGTCCATGATGATATACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 613
DB 800 AGTCTGTGAGAAATTAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 859
QY 614 ATGAGAGCCGGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 673
DB 860 ATGAGAGCCGGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 916
QY 674 AGCTTGGGCGCTGAGTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 733
DB 917 AGCTGCGCTCTGTGAATCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 976
QY 734 AGGAGATGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 793
DB 977 GAGAGAGAGCTCCGACTATGAGAGATTAACAGAGAGAGAGAGAGAGAGAGAGAGAG 1036
QY 794 GCTTGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 853
DB 1037 TCTGTCTG-----AACTGAGACTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1088
QY 854 CCTGCCCTTGTCTCAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 913
DB 1089 CTCTGCTTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1136
QY 914 TGGGGTTCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 973
DB 1137 TGGGGTTCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1194
QY 974 TCCCGTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1033
DB 1195 CTGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1248
QY 1034 AAAAAAAAAAAAAA 1045
DB 1249 AAAAAAAAAAAAAA 1260

```

```

RESULT 10
LOCUS RNJ001184 1188 bp mRNA linear ROD 10-FEB-1999
DEFINITION Rattus norvegicus mRNA for 36 kDa phosphoprotein.
ACCESSION AJ001184
VERSION AJ001184.1 GI:3021582
KEYWORDS 36 kDa phosphoprotein.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1
REFERENCE
AUTHORS Weber, J.R., Orstavik, S., Torgersen, K.M., Danbolt, N.C., Berg, S.F.,
Ryan, J.C., Taeken, K., Imboden, J.B. and Vage, J.T.
TITLE Molecular cloning of the cDNA encoding pp36, a
tyrosine-phosphorylated adaptor protein selectively expressed by T
cells and natural killer cells
JOURNAL J. Exp. Med. 187 (7), 1157-1161 (1998)
MEDLINE 98197173
PUBMED 9529333
REFERENCE
2 (bases 1 to 1188)
ORSTAVIK, S.
DIRECT SUBMISSION
Submitted (08-JAN-1998) Orstavik S., Institute of medical
biochemistry, University of Oslo, pb 1112, Blindern, N-0317 Oslo,
NORWAY

```

FEATURES

source

Location/Qualifiers

1..1188

/organism="Rattus norvegicus"

/mol_type="mRNA"

/strain="PVG"

/db_xref="taxon:10116"

/clone="151"

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 24475924. Location/Qualifiers

FEATURES

source

1. 1232

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="MGC:13859 IMAGE:419183"

/issue_type="Salivary gland, 10 week old female mouse"

/clone_id="NCI CGAP_SG2"

/lab_host="DH10B"

/note="Vector: pCMV-Sport6"

1. 1232

/gene="Lat"

/db_xref="LOCUSID:16797"

/db_xref="MGI:1342293"

269. 997

/gene="Lat"

/codon_start=1

/product="linker for activation of T cells"

/protein_id="AAH13337.1"

/db_xref="GI:15426469"

/db_xref="MGI:1342293"

/translation="MEADALSPVGLGLLLPLVTLAALCYRCRELPVSYDSTES

LYPSILIKPQITVPRPAVSPVPLSPPLROPDLPLPSPOILGSHMPSQO

NDDNANSVASYENOPRACKNVADBEDEDYDNGYLVLEDDSPANPVSAPVBNP

DIQDSAFVSECEIDVNPVSEBSASASLDSREYVNSPEQOPVTRALIASVNSEV

EEGEEBEGVDSBAPDYEINLEIN"

ORIGIN

Query Match 37.3%; Score 395.4; DB 10; Length 1232;

Best Local Similarity 67.5%; Pred. No. 7.5e-89;

Matches 693; Conservative 0; Mismatches 291; Indels 43; Gaps 8;

26 GTGACAGCAGCTGCTCCAGAGCTCCCTCAGATGAGAGAGGCGATCTGTCCTGCG 85
237 GTGACAGCAGGCGCTGTGAGAGGCTCTCAGATGAGAGAGGCGCTTGAAGCCGGTGG 296
86 TGTGGGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 142
297 GGTGGGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 356
143 GGCACAGACTGCAGAGCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 202
357 GCGTGTGAGTGCAGAGCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 416
203 TCCAGTTCAAGCGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 256
417 TCTCATCAAGCAGCAGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 476
257 TCACCTCTCAAGCAGCAGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 316
477 TCACCTCTCTCAAGCAGCAGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 536
317 CCCTTGGGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 376
537 CCCTTGGGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 596
377 TGGAGAGTACAG 436
597 TGGAGAGTACAG 656
437 ATCAACAGCCT--AGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 493
657 AGGACTATCCCAAGCGCTACTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 716
494 CTGCCCCATCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 553
717 TTGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 776
554 AGTCCATTGATGATTACGTAAAGCTTCCGAGAGCGGGAGAGAGAGAGAGAGAGAG 613

Db 777 AGTCGTGTAAGATTACGTAAAGTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 836
Qy 614 ATGACAGCCGGAGAGATGTGAATGTGTCTCCAGAGAGAGAGAGAGAGAGAGAGAG 673
Db 837 ATGAGAGCCGGAGAGATGTGAATGTGTCTCCAGAGAGAGAGAGAGAGAGAGAGAG 893
Qy 674 AGCTTCCGCGCTGAGTGTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 733
Db 894 AGCTGCTCTGTGAATCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 953
Qy 734 AGGAGATGTGAG 793
Db 954 GAGAGAGAGCTCCGAGCTAGAGAGATCTACAGAGAGAGAGAGAGAGAGAGAGAG 1013
Qy 794 GCTTGGCTTGGAG 853
Db 1014 TCTGTCTCTG-----AAACTGAGACTTGTGCTGCTGCTGCTGCTGCTGCTGCT 1065
Qy 854 CCGGCGCTTGTCTCAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 913
Db 1066 CTCTGCTTGTCCAG 1113
Qy 914 TGGGGTTCGGCTGTGTCTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 973
Db 1114 TGGGGTTCAGTGTGTGTCTCC--AAATCTGTACCTTGTGAATAAGCTGAGAGAT 1171
Qy 974 TGGCTTGGCGCCAG 1033
Db 1172 CTGTGTCCAG 1225
Qy 1034 AAAAAA 1040
Db 1226 AAAAAA 1232

RESULT 12
AC112166/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AC112166 131839 bp DNA linear PRI 23-AUG-2002
Homo sapiens chromosome 16 clone CTB-134H23, complete sequence.
AC112166
AC112166.3 GI:22450572
HTG.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
DOE Joint Genome Institute.
Sequencing of Human Chromosome 16
Unpublished
2 (bases 1 to 131839)
DOE Joint Genome Institute.
Direct Submission
Submitted (20-FEB-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 131839)
DOE Joint Genome Institute.
Direct Submission
Submitted (15-MAR-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
4 (bases 1 to 131839)
DOE Joint Genome Institute.
Direct Submission
Submitted (23-AUG-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Aug 23, 2002 this sequence version replaced gi:1948235.
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.

Base-by-base quality values are not generally visible from the Genbank flat file format but are available as part of this entry's ASN.1 file.

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the Genbank flat file format but are available as part of this entry's ASN.1 file.

FEATURES

Location/Qualifiers
1. .131839
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="CTB-134H23"

ORIGIN

Query Match 25.0%; Score 265; DB 9; Length 131839;
Best Local Similarity 98.2%; Pred. No. 8.9e-56;
Matches 268; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 757 TGAGGGGCTGTGAGGGCCGAGTCTGTCTGTGACACGAGCTTGGCTGGAGGAGGCTGAGCTG 816

DB 110544 TGTGTTTCAGTGGAGGGCCGAGTCTGTCTGTGAAACAGGCTTGGCTGGAGGAGGCTGAGCTG 110485

QY 817 GGCAGCTGGAAGTGGCTCTGGGGTCTTCACATGGCGTCTCCCTGTCTCCAGCCTGACA 876

DB 110484 GGCAGCTGGAAGTGGCTCTGGGGTCTTCACATGGCGTCTCCCTGTCTCCAGCCTGACA 110425

QY 877 ACAGCTGGAAGAAATCCCCCGTAATTTATCACTTTGGGGTTTCGGCTGTGTCCCCCG 936

DB 110424 ACAGCTGGAAGAAATCCCCCGTAATTTATCACTTTGGGGTTTCGGCTGTGTCCCCCG 110365

QY 937 AACGCTTGACACTTCTGACGAGCCTGAGATGACCTGCGGCCCGCCAGCCCTACTCT 996

DB 110364 AACGCTTGACACTTCTGACGAGCCTGAGATGACCTGCGGCCCGCCAGCCCTACTCT 110305

QY 997 GTGTATAGATTAAGGCTGCTGTGTCTGTG 1029

DB 110304 GTGTATAGATTAAGGCTGCTGTGTCTGTG 110272

RESULT 13

AC109460 163424 bp DNA linear PRI 12-OCT-2002
LOCUS Homo sapiens chromosome 16 clone RPl1-264B17, complete sequence.
DEFINITION
ACCESSION AC109460 GI:23915447
VERSION
KEYWORDS
SOURCE
ORGANISM

Homo sapiens (human)
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

DOE Joint Genome Institute, Stanford Human Genome Center and Los
Alamos National Laboratory.

Direct Submission

Unpublished

2 (bases 1 to 163424)

DOE Joint Genome Institute.

Direct Submission

Submitted (04-FEB-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

3 (bases 1 to 163424)

DOE Joint Genome Institute.

Direct Submission

Submitted (03-SEP-2002) Production Sequencing Facility, DOE Joint
Genome Institute

Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
4 (bases 1 to 163424)
DOE Joint Genome Institute, Stanford Human Genome Center and Los
Alamos National Laboratory.

Direct Submission

Submitted (12-OCT-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Oct 12, 2002 this sequence version replaced gi:22657513.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov

Finishing Completed at Stanford Human Genome Center and Los Alamos
National Laboratory
www.sngc.stanford.edu
Quality: Phrap Quality >=40 99.8% of Sequence;
Estimated Total Number of Errors is 0.3.

FEATURES

Location/Qualifiers
1. .163424
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RPl1-264B17"

ORIGIN

Query Match 25.0%; Score 265; DB 9; Length 163424;
Best Local Similarity 98.2%; Pred. No. 9e-56;
Matches 268; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 757 TGAGGGGCTGTGAGGGCCGAGTCTGTCTGTGAAACAGGCTTGGCTGGAGGAGGCTGAGCTG 816

DB 122732 TGTGTTTCAGTGGAGGGCCGAGTCTGTCTGTGAAACAGGCTTGGCTGGAGGAGGCTGAGCTG 122791

QY 817 GGCAGCTGGAAGTGGCTCTGGGGTCTTCACATGGCGTCTCCCTGTCTCCAGCCTGACA 876

DB 122792 GGCAGCTGGAAGTGGCTCTGGGGTCTTCACATGGCGTCTCCCTGTCTCCAGCCTGACA 122851

QY 877 ACAGCTGGAAGAAATCCCCCGTAATTTATCACTTTGGGGTTTCGGCTGTGTCCCCCG 936

DB 122852 ACAGCTGGAAGAAATCCCCCGTAATTTATCACTTTGGGGTTTCGGCTGTGTCCCCCG 122911

QY 937 AACGCTTGACACTTCTGACGAGCCTGAGATGACCTGCGGCCCGCCAGCCCTACTCT 996

DB 122912 AACGCTTGACACTTCTGACGAGCCTGAGATGACCTGCGGCCCGCCAGCCCTACTCT 122971

QY 997 GTGTATAGATTAAGGCTGCTGTGTCTGTG 1029

DB 122972 GTGTATAGATTAAGGCTGCTGTGTCTGTG 123004

RESULT 14

AC138925 187788 bp DNA linear HTG 21-JAN-2003
LOCUS Homo sapiens chromosome 16 clone RPl1-664F22, WORKING DRAFT
DEFINITION
ACCESSION AC138925
VERSION
KEYWORDS
SOURCE
ORGANISM

Homo sapiens (human)
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

DOE Joint Genome Institute.

Sequencing of Human Chromosome 16

Unpublished

2 (bases 1 to 187788)

DOE Joint Genome Institute.

Direct Submission

Submitted (21-JAN-2003) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

-----Genome Center

Center: Joint Genome Institute

Center Code: JGI

Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 1549622
Center Clone name: RPCT-11_664F22

Summary Statistics
Consensus quality: 186934 bases at least Q40
Consensus quality: 187077 bases at least Q30
Consensus quality: 187284 bases at least Q20
Estimated insert size: 180000; agarose-fp estimation
Estimated insert size: 187688; sum-of-ctnigs estimation
Quality coverage: 11.48 in Q20 bases; agarose-fp estimation
Quality coverage: 11.01 in Q20 bases; sum-of-ctnigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 34623: contig of 34623 bp in length
* 34624 34723: gap of unknown length
* 34724 187788: contig of 153065 bp in length.
Location/Qualifiers
1.187788
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-664F22"
/clone_1ib="RPCT human BAC library 11"

ORIGIN

Query Match 25.0%; Score 265; DB 2; Length 187788;
Best Local Similarity 98.2%; Pred. No. 9.1e-56;
Matches 268; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 757 TGAAGGCTGTGAGAGCCGAGTGTCTCTGGAACCAAGCTTGCGAGGCTGAGCTG 816
DB 111946 TGTGTTTCAGTGGAGCCGAGTGTCTCTGGAACCAAGCTTGCGAGGCTGAGCTG 111887

QY 817 GGCAGCTGGAAGTGGCTGCGGGGCTCTCAATGAGGCTGCTGCTGCTCCAGCCGAGA 876
DB 111886 GGCAGCTGGAAGTGGCTGCGGGGCTCTCAATGAGGCTGCTGCTGCTCCAGCCGAGA 111827

QY 877 ACAGCTGGAAGTGGCTGCGGGGCTCTCAATGAGGCTGCTGCTGCTCCAGCCGAGA 936
DB 111826 ACAGCTGGAAGTGGCTGCGGGGCTCTCAATGAGGCTGCTGCTGCTCCAGCCGAGA 111767

QY 937 AACGCTGCGACCTTCTGAGCAGCCTGGAATGACCTGCGCCGAGCCCTACTCT 996
DB 111766 AACGCTGCGACCTTCTGAGCAGCCTGGAATGACCTGCGCCGAGCCCTACTCT 111707

QY 997 GTGTAATGAATTAAGCCTGCGTGTCTGTG 1029
DB 111706 GTGTAATGAATTAAGCCTGCGTGTCTGTG 111674

RESULT 15
AC139247
DEFINITION Homo sapiens chromosome 16 clone CTD-2553D4, WORKING DRAFT
AC139247
SEQUENCE, 18 unordered pieces.
AC139247
AC139247.1 GI:28008725
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 219646)

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL
COMMENT

DOE Joint Genome Institute.
Sequencing of Human Chromosome 16
Unpublished
2 (bases 1 to 219646)
DOE Joint Genome Institute.
Direct Submission
Submitted (29-JAN-2003) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 810409
Center Clone name: CTD-2553D4

Summary Statistics
Consensus quality: 211019 bases at least Q40
Consensus quality: 212006 bases at least Q30
Consensus quality: 212932 bases at least Q20
Estimated insert size: 175000; agarose-fp estimation
Estimated insert size: 217946; sum-of-ctnigs estimation
Quality coverage: 31.21 in Q20 bases; agarose-fp estimation
Quality coverage: 25.06 in Q20 bases; sum-of-ctnigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1100: contig of 1100 bp in length
* 1101 1200: gap of unknown length
* 1201 2326: contig of 1126 bp in length
* 2327 2426: gap of unknown length
* 2427 3448: contig of 1022 bp in length
* 3449 3548: gap of unknown length
* 3549 4793: contig of 1245 bp in length
* 4794 4893: gap of unknown length
* 4894 5992: contig of 1099 bp in length
* 5993 6092: gap of unknown length
* 6093 8163: contig of 2071 bp in length
* 8164 8263: gap of unknown length
* 8264 9387: contig of 1124 bp in length
* 9388 9487: gap of unknown length
* 9488 10873: contig of 1386 bp in length
* 10874 10973: gap of unknown length
* 10974 12188: contig of 1215 bp in length
* 12189 12288: gap of unknown length
* 12289 13292: contig of 1004 bp in length
* 13293 13392: gap of unknown length
* 13393 14698: contig of 1306 bp in length
* 14699 14798: gap of unknown length
* 14799 16556: contig of 1758 bp in length
* 16557 16656: gap of unknown length
* 16657 18479: contig of 1823 bp in length
* 18480 18579: gap of unknown length
* 18580 22438: contig of 3859 bp in length
* 22439 22539: gap of unknown length
* 22539 40064: contig of 17526 bp in length
* 40065 40165: gap of unknown length
* 40165 53730: contig of 13566 bp in length
* 53731 53831: gap of unknown length
* 53831 89191: contig of 35361 bp in length
* 89192 89291: gap of unknown length
* 89292 219646: contig of 130355 bp in length.
Location/Qualifiers
1.219646
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

/chromosome="16"
/clone="CTD-2553D4"
/clone_lib="Caltech human BAC library D"

ORIGIN

Query Match 25.0%; Score 265; DB 2; Length 219646;
Best Local Similarity 98.2%; Pred. No. 9.2e-56;
Matches 268; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 757 TGAGGGCCTGTGAGAGCCGAGCTGTCTCTGGAACGAGCTTGCTGGGAGCGGTGAGCTG 816
|||
Db 130373 TGTGTTTCAGTGAAGGCCGAGCTGTCTCTGGAACGAGCTTGCTGGGAGCGGTGAGCTG 130432

QY 817 GGCAGCTGGAAGTGGCTCTGGGGTCTCTCAATGGCGTCTGCCCTTGCTCCAGCTGACA 876
|||
Db 130433 GGCAGCTGGAAGTGGCTCTGGGGTCTCTCAATGGCGTCTGCCCTTGCTCCAGCTGACA 130492

QY 877 ACAGCCTGAGAAATCCCCCGTAATTAATCACTTTGGGGTTGCGGCTGTGTCCCCG 936
|||
Db 130493 ACAGCCTGAGAAATCCCCCGTAATTAATCACTTTGGGGTTGCGGCTGTGTCCCCG 130552

QY 937 AACGCTGTGACCTTCTGAGCGCAGCTGAGAAATGACCTGCGCCCGCCAGCCCTACTCT 996
|||
Db 130553 AACGCTGTGACCTTCTGAGCGCAGCTGAGAAATGACCTGCGCCCGCCAGCCCTACTCT 130612

QY 997 GTGTATAGATTAAGGCTGCGTGTGTGTG 1029
|||
Db 130613 GTGTATAGATTAAGGCTGCGTGTGTGTG 130645

Search completed: November 22, 2004, 10:12:47
Job time : 6516 secs

CC substrates and intracellular ligands or between these ligands and other
 CC members of the pathway, including identification of downstream signalling
 CC proteins, particularly in immune system cells. These modulators are
 CC potentially useful as drugs and diagnostic agents, particularly for
 CC diseases that involve undesirable cell proliferation, differentiation,
 CC growth or T cell energy, e.g. neoplasia, inflammation, hypersensitivity/
 CC allergy, microbial infection, metabolic, genetic or autoimmune diseases,
 CC graft rejection. LAT is also used to generate specific antibodies, used
 CC for detection of LAT. Nucleic acid that encodes LAT, or its fragments,
 CC are used to identify homologous sequences in other species, to detect the
 CC LAT gene and as sources of antisense therapeutics. Modulators of LAT are
 CC potentially more specific and less toxic than known immunosuppressants
 CC such as cyclosporin. The present sequence represents the nucleotide
 CC sequence of human LAT

SO Sequence 1059 BP; 227 A; 330 C; 301 G; 201 T; 0 U; 0 Other;

Query Match 98.9%; Score 1048; DB 2; Length 1059;

Best Local Similarity 99.9%; Pred. No. 2,2e-233;

Matches 1059; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GACTCTGCCCTTGAGGGGCTTGAAGGCTGACGACCCCTGCTCCAGCTCCCTGCAGATG 60
 DB 1 GACTCTGCCCTTGAGGGGCTTGAAGGCTGACGACCCCTGCTCCAGCTCCCTGCAGATG 60
 QY 61 GAGGAGGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
 DB 61 GAGGAGGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 119
 QY 121 TTGATGCACTGTGTGTGCACTGCTGCAAGACTGCTGCAAGCTCTTACAGACAGATCTCTCA 180
 DB 120 TTGATGCACTGTGTGTGCACTGCTGCAAGACTGCTGCAAGCTCTTACAGACAGATCTCTCA 179
 QY 181 GATGTTTGTATCCAAAGGGGATCAAGTTTAAAGGCTTCAAGGTTGCCCCCTTGCCCA 240
 DB 180 GATGTTTGTATCCAAAGGGGATCAAGTTTAAAGGCTTCAAGGTTGCCCCCTTGCCCA 239
 QY 241 CTTGCTACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
 DB 240 CTTGCTACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 299
 QY 301 CCAAGATCCCGGAGCCCTTGGGGGCTCCCAAGGAGCCGATTTCCCGGGGAGATTCT 360
 DB 300 CCAAGATCCCGGAGCCCTTGGGGGCTCCCAAGGAGCCGATTTCCCGGGGAGATTCT 359
 QY 361 GATGCTGCCCAAGTGTGCGAGCTACGAGAACGAGAACGAGCTGTGAGGATGCAAGT 420
 DB 360 GATGCTGCCCAAGTGTGCGAGCTACGAGAACGAGAACGAGCTGTGAGGATGCAAGT 419
 QY 421 GAGGATGAGGAGCACTATCAACAACCAAGGCTACCTGCTGCTGCTGCTGCTGCTGCTG 480
 DB 420 GAGGATGAGGAGCACTATCAACAACCAAGGCTACCTGCTGCTGCTGCTGCTGCTGCTG 479
 QY 481 GCCACTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 DB 480 GCCACTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 539
 QY 541 GCTTCTCCATGAGTGCATTGATGATTAAGTAAAGTTCGGAGAGCGGGGAGAGGCA 600
 DB 540 GCTTCTCCATGAGTGCATTGATGATTAAGTAAAGTTCGGAGAGCGGGGAGAGGCA 599
 QY 601 GAAGCGTCTGATGAGGAGCGGGAGATGATGATGATGATGATGATGATGATGATGATG 660
 DB 600 GAAGCGTCTGATGAGGAGCGGGAGATGATGATGATGATGATGATGATGATGATGATG 659
 QY 661 GCGGCTAAGACTGAGCGCTGCGCTTGAAGTTCCAGAGAGGAGAGAGAGAGAGAG 720
 DB 660 GCGGCTAAGACTGAGCGCTGCGCTTGAAGTTCCAGAGAGGAGAGAGAGAGAGAG 719
 QY 721 GGGGCTCAAGTTACAGAAATCTTCAGAGAGTGAAGTGAAGGCTGTGAGAGGCGGAGTCT 780
 DB 720 GGGGCTCAAGTTACAGAAATCTTCAGAGAGTGAAGTGAAGGCTGTGAGAGGCGGAGTCT 779

QY 781 GTCTGAGACAGAGCTTGCCTTGAGAGCGCTGAGCTGGGAGCTGGAAGTGGCTTGAGGT 840
 DB 780 GTCTGAGACAGAGCTTGCCTTGAGAGCGCTGAGCTGGGAGCTGGAAGTGGCTTGAGGT 839
 QY 841 CTTCACTGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 DB 840 CTTCACTGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 899
 QY 901 CTTATATATCACTTGGGGGCTTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
 DB 900 CTTATATATCACTTGGGGGCTTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 959
 QY 961 CTTGAGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 DB 960 CTTGAGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1019
 QY 1021 GTGCTGTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1060
 DB 1020 GTGCTGTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1059

RESULT 2

AA206221 standard; DNA; 1224 BP.

AA206221;

30-SEP-1999 (first entry)

Human secreted protein gene No. 3.

Human; secreted protein; fusion protein; gene therapy; protein therapy;
 diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 developmental abnormality; fetal deficiency; blood; allergy; renal; ds;
 immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

Homo sapiens.

MO9935158-A1.

15-JUL-1999.

06-JAN-1999; 99MO-US000108.

07-JAN-1998; 98US-0070657P.

07-JAN-1998; 98US-0070658P.

07-JAN-1998; 98US-0070658P.

07-JAN-1998; 98US-0070704P.

(HUMA-) HUMAN GENOME SCI INC.

Ruben SM, Soppet DR, Ebner R, Lafleur DW, Ni J, Brewer LA;

P-Olsen HS, Duan RD, Rosen CA;

WPI; 1999-444190/37.

P-PSDB; AAY38388, AAY38432, AAY38433, AAY38434.

New isolated human genes and the secreted polypeptides they encode.

Claim 1, Page 151; 227bp; English.

This sequence represents a nucleic acid molecule which encodes a secreted
 human protein. The gene number is given in the descriptor line. The gene
 can be used to generate fusion proteins by linking to the gene to a human
 immunoglobulin Fc portion (e.g. AA206210) for increasing the stability of
 the fused protein as compared to the human protein only. The invention
 relates to 36 novel genes and their fragments (nucleic acid sequences:
 AA206219-206263; amino acid sequences AAY38386-Y38498) which are useful
 for preventing, treating or ameliorating medical conditions e.g. by

CC protein:nerve gene therapy. Also, pathological conditions can be diagnosed
CC by determining the amount of the new polypeptides in a sample or by
CC determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 35 polynucleotides, based on
CC which tissues they are most highly expressed in (see AM206219 for
CC described uses)
CC

SQ Sequence 1224 BP; 243 A; 404 C; 346 G; 230 T; 0 U; 1 Other;

Query Match	89.5%	Score 949.2;	DB 2;	Length 1224;
Best Local Similarity	98.9%	Pred. No. 2e-210;		
Matches 1039; Conservative	0;	Mismatches 4;	Indels 8;	Gaps 8

QY	1	GACTCTGCGCTTGAGGGGCGCTAAGGGGATGACAGCACCTGCTCCAGACTCCCTCGAAGT	60
Db	181	GACTCTGCGCTTGAGGGGCGCTAAGGGGATGACAGCACCTGCTCCAGACTCCCTCGAAGT	240
QY	61	GAGAGGCGCATCTGCTGCTCCCTGCTGCGTGGGGCTCTGCTGCTGGCCATCTGGACATG	120
Db	241	GAGAGGCGCATCTGCTGCTCCCTGCGTGGGGGCTCTGCTGCTGGCCATCTGGACATG	300
QY	121	TTGATGGCACTGTGTGTGACCTGGCACAAGCTGCCAGGCTCTTACAGACAGCAATCTCA	180
Db	301	TTGATGGCACTGTGTGTGACCTGGCACAAGCTGCCAGGCTCTTACAGACAGCAATCTCA	360
QY	181	GATAGTTTGTATCC-AAGGGGCGATCTCAGTTGAAACGGGCTGACAAGGTGTGCCCCCTGGCC	239
Db	361	GATAGTTTGTATCCAAAGGGGCGATCAGTTGAAACGGGCTGACAAGGTGTGCCCCCTGGCC	420
QY	240	ACCTGCGTACCCCACTGTCACTCTCTACCCACCCCTGAGCCAGCAAGCTGCTCCCAT	299
Db	421	ACCTGCGTACCCCACTGTCACTCTCTACCCACCCCTGAGCCAGCAAGCTGCTCCCAT	480
QY	300	CCCAAGATCCCGGAGCGCCCTTGGGGGGCTCCCAACGGACGCCATCTTCCCGGCGGGATTC	359
Db	481	CCCAAGATCCCGGAGCGCCCTTGGGGGGCTCCCAACGGACGCCATCTTCCCGGCGGGATTC	540
QY	360	TGATGTGTCCAAAGTGTGTGCGAGCTACGAGAACGAGGAACCAAGCTGTGAGATGACAGA	419
Db	541	TGATGTGTGTCCAAAGTGTGTGCGAGCTACGAGAACCAAGCTGTGAGATGACAGA	599
QY	420	TGAGAGTAGAGGACGACTATCAACAACCGAGGCTACCTGGTGTGCTTCTTGACAGACCCC	479
Db	600	TGA-GATGAGGACGACTATCAACAACCGAGGCTACCTGGTGTGCTTCTTGACAGACCCC	658
QY	480	GCGCACTAGACAGCTGTGCCCATCAGCTCTCTGCACTCAGCAACCCCTGGCATCTCGAAG	539
Db	659	GCGCACTAGACAGCTGTGTGCCCATCAGCTCTCTGCACTCAGCAACCCCTGGCATCTCGAAG	717
QY	540	TGCTTTCTCCATGAGATCATTGATGATTAAGTGAAGTTCCGAGAGCGGGGAGAGCGC	599
Db	718	TGCTTTCTCCATGAGATCATTGATGATTAAGTGAAGTTCCGAGAGCGGGGAGAGCGC	777
QY	600	AGAAGCGTCTGTGATGGCAGCCGGAGATATGTGAATGTGTCCAGAAATGCTCATCTGG	659
Db	778	AGAAGCGTCTGTGATGGCAGCCGGAGATATGTGAATGTGTGTCCAGAAATGCTCATCTGG	837
QY	660	AGCGGCTTACAGCTGAGCGCTCGCGCTCTGAGTTCCAGAGGCGAAGAGATGTGAGAGAA	719
Db	838	AGCGGCTTACAGCTGAGCGCTCGCGCTCTGAGTTCCAGAGGCGAAGAGATGTGAGAGAA	897
QY	720	GGGGGCTTCAGATTACGAGAACTTCGACAGACTGAACTGAAGGGGCTGTGAGAGGCGGAGTTC	779
Db	898	GGGGGCTTCAGATTACGAGAACTTCGACAGACTGAACTGAAGGGGCTGTGAGAGGCGGAGTTC	957
QY	780	TGTCCTTGAAACAGAGCTTGCTGTGGAGCGGCTGAGCTGGGACAGCTGGAAGTGGCTTGGGG	839
Db	958	TGTCCTTGAAACAGAGCTTGCTGTGGAGCGGCTGAGCTGGGACAGCTGGAAGTGGCTTGGGG	1012
QY	840	TTCCTCACATGGCGTCTGCGCCTTGCTGCTCAAGCTGACACAGCTGTAGAAATTCCTCCCGTA	899
Db	1018	TTCCTCACATGGCGTCTGCGCCTTGCTGCTCAAGCTGACACAGCTGTAGAAATTCCTCCCGTA	1077

Qy	900	ACTATATATACACTTTGGGGTTGGCCGCTGTCGCCGAAGCTGTGACCTCTGACGCA	959
Db	1078	ACTATTATATACATTTGGGGTTGGCCGCTGTG-CCCCCAAGCGCTGTG-ACCTTTGACGCA	1135
Qy	960	GCCTGAGATGACCTGGCCCTGGCCCCCAGCCCTACTCTGTGTAATAGATATAAGGCTGCG	1019
Db	1136	GCCTGAGATGACCTGTG-CTGGCCCCCAGCCCTACTCTGTG-GTAATACATATAAGGCTGCG	1193
Qy	1020	TGTGCTGTGGAAAAAAAAAAAAAAAAAAAAA	1050
Db	1194	TGTGCTGTGMAAAAAAAAAAAAAAAAAAAAAA	1224

RESULT 3
ADQ22892
ID ADQ22892 standard; DNA; 1488 BP

AC	ADQ22892;
XX	
DT	26-AUG-2004 (first entry)
XX	

DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 5712.

KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW ds.

OS Homo sapiens.

PN WO2004048938-A2.

PD 10-JUN-2004.

PF 26-NOV-2003; 2003WO-US038193.

PR 26-NOV-2002; 2002US-0429739P.

PA (PROT-) PROTEIN DESIGN LABS INC.

PI Aziz N, Ginsburg WM, Zlotnik A;

DR WPI; 2004-441208/41.

PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.

PS Example 2; SEQ ID NO 5712; 210pp; English.

CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual,
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cyrostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
XX Sequence 1488 BP; 294 A; 443 C; 432 G; 319 T; 0 U; 0 Other;
XQ

Query Match	87.9%	Score 932;	DB 12;	Length 1488;
Best Local Similarity	92.2%	Pred. No. 2.1e-206;		
Matches 1029; Conservative	0;	Mismatches 0;	Indels 87;	Gaps 1.

[illegible]

QY	6	PAGAGAGGCATCTGAGTCCCGGAGTGGGGGCTCTGCTGCTGCCATCTGGGCAATG	120
Db	82	GAGAGAGGCATCTGGTCCCTCGGTGCTGGGGCTCTGCTGCTGCCATCTGGGCAATG	141
QY	121	TTGATGCACTGTGTGTGCACTGCCACAGACTGCCAGGCTCTTAAGACAGCATCTCA	180
Db	142	TTGATGCACTGTGTGTGCACTGCCACAGACTGCCAGGCTCTTAAGACAGCATCTCA	201
QY	181	GATAGTTTGTATCCAAAGGGGCACTCGATTCAAAGGCTTCAAGGTTGGCCCCCTGGCCA	240
Db	202	GATAGTTTGTATCCAAAGGGGCACTCGATTCAAAGGCTTCAAGGTTGGCCCCCTGGCCA	261
QY	241	CCTGCGCTACCCCACTGTCACTCTCAACCAACCCCTGAGCGACCCAGACTCTGCTCCCATC	300
Db	262	CCTGCGCTACCCCACTGTCACTCTCAACCAACCCCTGAGCGACCCAGACTCTGCTCCCATC	321
QY	301	CCAGATCCCGGACAGCCCTTGGGGGCTTCCACCGGACGCATCTTCCCGGCGGAAATTCT	360
Db	322	CCAGATCCCGGACAGCCCTTGGGGGCTTCCACCGGACGCATCTTCCCGGCGGAAATTCT	381
QY	361	GATGTTGCCAACAGTGTGGCGAGCTACGAGAACGAG-----	396
Db	382	GATGTTGCCAACAGTGTGGCGAGCTACGAGAACGAGGGTGGCTTGGGATCCGAGGTGCC	441
QY	397	-----	396
Db	442	CAGGCTGGGTGGGAGAGTCTGGGGGTCCGTCCTGGACTAGGCTGACCCCTGTGCTTACCC	501
QY	397	---GAAACAGCCTGTGAGAGATGCAATAGAGATGAGAGACGATTAATACAAACCAAGCTAC	453
Db	502	CCAGAACAGCCTGTGAGAGATGCAATAGAGATGAGAGACGATTAATACAAACCAAGCTAC	561
QY	454	CTGGTGTGTCTTCTTGACAGCACCCCGGCACTAGCACTGTGCCCATCAGCTCTCTGA	513
Db	562	CTGGTGTGTCTTCTTGACAGCACCCCGGCACTAGCACTGTGCCCATCAGCTCTCTGA	621
QY	514	CTCAGCACCCCTGGCATCCGAGACAGTGTCTTCTCCATGAGAGTCCAATTGATATTAAGTG	573
Db	622	CTCAGCACCCCTGGCATCCGAGACAGTGTCTTCTCCATGAGAGTCCAATTGATATTAAGTG	681
QY	574	AAACGTTCCGGAGAGGGGGGAGAGCCCAAGAAAGGCTCTGAGATGAGCACCCGGGAGATATGG	633
Db	682	AAACGTTCCGGAGAGGGGGGAGAGCCCAAGAAAGGCTCTGAGATGAGCACCCGGGAGATATGG	741
QY	634	AATGTGTCCCAAGGAATGTGCATCTGTGAGCGGCTAAGACTGAGCCTTGCCTGTGATTTCC	693
Db	742	AATGTGTCCCAAGGAATGTGCATCTGTGAGCGGCTAAGACTGAGCCTTGCCTGTGATTTCC	801
QY	694	CAGGAGGCAAGAGAAATGAGAGAAAGAGGGGGCTTCCAGATTACGAGAAATTGACAGAGCTG	753
Db	802	CAGGAGGCAAGAGAAATGAGAGAAAGAGGGGGCTTCCAGATTACGAGAAATTGACAGAGCTG	861
QY	754	AACTGAGGGGCTGTGAGAGGCGAGTCTGTCTTGGAACAAGGCTTGGCTGGAGCGGCTGAG	813
Db	862	AACTGAGGGGCTGTGAGAGGCGAGTCTGTCTTGGAACAAGGCTTGGCTGGAGCGGCTGAG	921
QY	814	CTGGGCACTGGAAGTGTGCTTGTGGGGTCTTCACTAGTGGCTGTGCCCTTGTCCACGCTTG	873
Db	922	CTGGGCACTGGAAGTGTGCTTGTGGGGTCTTCACTAGTGGCTGTGCCCTTGTCCACGCTTG	981
QY	874	ACAAAGGCTGAGAAATCCCCCGGTAATTATTAACCTTGGGGGTGGGCTGTGTGCC	933
Db	982	ACAAAGGCTGAGAAATCCCCCGGTAATTATTAACCTTGGGGGTGGGCTGTGTGCC	1041
QY	934	CCGAAGGCTGTGACCTTCTGACGAGCCTGGAATGACTGCTGCCCTGGCCCAAGCCTTAC	993
Db	1042	CCGAAGGCTGTGACCTTCTGACGAGCCTGGAATGACTGCTGCCCTGGCCCAAGCCTTAC	1101
QY	994	TCTGTGTATAGAAATTAAGAGCCTGCGTGTGTCTGTG 1029	
Db	1102	TCTGTGTATAGAAATTAAGAGCCTGCGTGTGTGTCTGTG 1137	

ID	AA	89073	standard; DNA; 1461 BP.
XX	AC	AA	89073;
XX	DT	14-SEP-1999	(first entry)
XX	XX	Human LAT alternative splice variant nucleotide sequence.	
XX	KW	LAT, tyrosine kinase; linker for activation of T cell; TCR; human;	
XX	KW	T-cell receptor; TCR signalling pathway; neoplasia; inflammation;	
XX	KW	hypersensitivity; allergy; microbial infection; genetic disease;	
XX	KX	autoimmune disease; graft rejection; modulator; variant; ss.	
XX	OS	Homo sapiens.	
XX	PN	MO9932627-A2.	
XX	PD	01-JUL-1999.	
XX	PF	23-DEC-1998; 98WO-US027400.	
XX	PR	23-DEC-1997; 97US-0068690P.	
XX	PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.	
XX	PI	Sameison LE, Zhang W;	
XX	DR	WPI, 1999-418926/35.	
XX	PT	Linker for activation of T cell protein used to, e.g. screen for	
XX	XX	modulators of T cell signaling.	
XX	PS	Disclosure; Fig 7B; 125p; English.	
XX	CC	The invention relates to a protein tyrosine kinase substrate LAT (linker	
XX	CC	for activation of T cells) protein. Modulation of interaction between LAT	
XX	CC	and the T-cell receptor (TCR) affects the TCR signalling pathway. LAT is	
XX	CC	a substrate for tyrosine kinases and becomes phosphorylated after TCR	
XX	CC	engagement, resulting in recruitment of other signalling molecules. LAT	
XX	CC	is used to identify and test (ant)agonists of tyrosine kinase signalling	
XX	CC	pathways, i.e. modulation of interaction between tyrosine kinase	
XX	CC	substrates and intracellular ligands or between these ligands and other	
XX	CC	members of the pathway, including identification of downstream signalling	
XX	CC	proteins, particularly in immune system cells. These modulators are	
XX	CC	potentially useful as drugs and diagnostic agents, particularly for	
XX	CC	diseases that involve undesirable cell proliferation, differentiation,	
XX	CC	growth or T cell anergy, e.g. neoplasia, inflammation, hypersensitivity/	
XX	CC	allergy, microbial infection, metabolic, genetic or autoimmune diseases,	
XX	CC	graft rejection. LAT is also used to generate specific antibodies, used	
XX	CC	for detection of LAT. Nucleic acid that encodes LAT, or its fragments,	
XX	CC	are used to identify homologous sequences in other species; to detect the	
XX	CC	LAT gene and as sources of antisense therapeutics. Modulators of LAT are	
XX	CC	potentially more specific and less toxic than known immunosuppressants	
XX	CC	such as cyclosporin. The present sequence represents the nucleotide	
XX	CC	sequence of an alternative splice variant of human LAT	
XX	CC	Sequence 1461 BP; 269 A; 443 C; 432 G; 317 T; 0 U; 0 Other;	
XX	CC	Query Match 86.9%; Score 921; DB 2; Length 1461;	
XX	CC	Best Local Similarity 92.1%; Pident. No. 7,4e-204;	
XX	CC	Matches 1029; Conservativity 0; Mismatches 0; Indels 88; Gaps 2;	
XX	CC	1 GAGCTGCGCCCTTGAAGGGGCTTGAAGGGGTGTCACACAGCCCTGTCGAGTCCCTGCAGATG 60	
XX	CC	22 GAGCTGCGCCCTTGAAGGGGCTTGAAGGGGTGTCACACAGCCCTGTCGAGTCCCTGCAGATG 81	
XX	CC	61 GAGAGGAGCATCTGATGTCCTCTGGGCTGTCGAGGGCTCTGCTGTCACATCTGGCCATG 120	
XX	CC	82 GAGAGGAGCATCTGATGTCCTCTGGGCTGTCGAGGGCTCTGCTGTCACATCTGGCCATG 141	
XX	CC	121 TTGATGCACTGTGTGTGCACTGCCACAGACTGCCAGGCTCCCTTACGACAGCATCTCTTA 180	

D	b	142	TTGATGGCACTGTGTGTGTGACCTGGCCACAGACTGCCAGGCTCTTACGACAGCAATCTCTCA	201
O	y	181	GATAGTTTGTATCCAAAGGGGCATCCAGTTCAAAAGGCTTCACACGGTTGGCCCTGGCCA	240
D	b	202	GATAGTTTGTATCCAAAGGGGCATCCAGTTCAAAAGGCTTCACACGGTTGGCCCTGGCCA	261
O	y	241	CCGTGCTTACCCCACTGTGCACCTTCCCTTACCCACCCCTTACGACGACAGCACTGTCTCCCATC	300
D	b	262	CTGTGCTTACCCCACTGTGCACCTTCCCTTACCCACCCCTTACGACGACAGCACTGTCTCCCATC	321
O	y	301	CCAAATCCCCCAGGCCCCCTTGGGGGCTCCACCCGACGCGCATC-TTCCCGGCGGATTC	359
D	b	322	CCAAATCCCCCAGGCCCCCTTGGGGGCTCCACCCGACGCGCATCTTTCCCGGCGGATTC	381
O	y	360	TGATGTGTCCAAACAGTGTGGCGAAGCTTACGAGAACGAG-----	396
D	b	382	TGATGTGTCCAAACAGTGTGGCGAAGCTTACGAGAACGAGGTGCGTGTGGATCCGAGGTGC	441
O	y	397	-----	396
D	b	442	CCAGGCTGGGTGGGAGATCTGGGGTCCGTCTTGACCTAAGCTGACCCCTGTGTCTTACC	501
O	y	397	----GAAACAGGCTGTGAGAGATGACAGATAGAGATGAGAGCACTATACAAACCAAGCTA	452
D	b	502	CCCAAGAACAGGCTGTGAGAGATGACAGATAGAGATGAGAGCACTATACAAACCAAGCTA	561
O	y	453	CCGTGTGTGTCTTCTGTGACAGACCCCGGCTCATACGCTGCTGCCCCATCACTCTGCG	512
D	b	562	CCGTGTGTGTGTCTTCTGTGACAGACCCCGGCTCATACGCTGCTGCCCCATCACTCTGCG	621
O	y	513	ACTCAGCACCCCTGGCATTCGAGACAGTGTCTTCTCATGTAGAGTCCATGTGATTAACGT	572
D	b	622	ACTCAGCACCCCTGGCATTCGAGACAGTGTCTTCTCATGTAGAGTCCATGTGATTAACGT	681
O	y	573	GAAAGCTTCCGGAGAGCGGGGAGAGCGGCACAAGGTCTCTGAGATGGGACCGGGAGATGT	632
D	b	682	GAAAGCTTCCGGAGAGCGGGGAGAGCGGCACAAGGTCTCTGAGATGGGACCGGGAGATGT	741
O	y	633	GAATGTGTCCCAAGAACTGCATCTTGAGACGGCTTAAGACTGAGCTGCGCCCTGAGTTC	692
D	b	742	GAATGTGTCCCAAGAACTGCATCTTGAGACGGCTTAAGACTGAGCTGCGCCCTGAGTTC	801
O	y	693	CCAGAGGCGAGAGAAGTGTGAGAGAAAGAGGGGCTCTCAGATTACGAGAACTTGCAGAGCT	752
D	b	802	CCAGAGGCGAGAGAAGTGTGAGAGAAAGAGGGGCTCTCAGATTACGAGAACTTGCAGAGCT	861
O	y	753	GAACTGAGAGGCTGTGTGAGAGCCGAGTCTGTCTTGAAACAGAGCTTGTGCTGGACGCTGA	812
D	b	862	GAACTGAGAGGCTGTGTGAGAGCCGAGTCTGTCTTGAAACAGAGCTTGTGCTGGACGCTGA	921
O	y	813	GCTGGGCAAGCTGGAAGTGGCTCTGGGGGTCTCAATGAGGCTCTGCGCTTGCTCCAGCCT	872
D	b	922	GCTGGGCAAGCTGGAAGTGGCTCTGGGGGTCTCAATGAGGCTCTGCGCTTGCTCCAGCCT	981
O	y	873	GACAAACAGCTGAGAAATCCCCCGTAATTATACCTTGTGGGGTGTGGCTGTGTCC	932
D	b	982	GACAAACAGCTGAGAAATCCCCCGTAATTATACCTTGTGGGGTGTGGCTGTGTGTCC	104
O	y	933	CCCGAAGCTGTGCACCTTCTGACGCAAGCTTGAAGAAATGACCTTGCCTGGCCCCAGCCTTA	992
D	b	1042	CCCGAAGCTGTGCACCTTCTGACGCAAGCTTGAAGAAATGACCTTGCCTGGCCCCAGCCTTA	110
O	y	993	CTCTGTGTAAATGAATTAAGGCGTGTGTGTGTG 1029	
D	b	1102	CTCTGTGTAAATGAATTAAGGCGTGTGTGTGTG 1138	

XX	26-AUG-2004	(first entry)	
XX	Human soft tissue sarcoma-upregulated DNA - SEQ ID 1195.		
XX	soft tissue sarcoma; cytoskeletal; gene therapy; vaccine; screening; human;		
XX	ds.		
XX	Homo sapiens.		
OS	MO2004048938-A2.		
PN	10-JUN-2004.		
XX	26-NOV-2003; 2003WO-US038193.		
XX	26-NOV-2002; 2002US-0429739P.		
XX	(PROT-) PROTEIN DESIGN LABS INC.		
PA	Aziz N, Gineburg WM, Zlotnick A;		
PI	WPI; 2004-441208/41.		
XX	Early detection of soft tissue sarcoma comprises determining expression		
XX	of a gene in a first soft tissue sample and a normal soft tissue sample		
XX	and comparing the gene expression, also useful in treating soft tissue		
XX	sarcoma.		
PT	Example 2; SEQ ID NO 1195; 210pp; English.		
PS	The invention relates to a novel method for detecting soft tissue sarcoma		
XX	which comprises obtaining a first soft tissue sample from an individual		
XX	and a normal soft tissue sample from the same or different individual,		
XX	determining the expression of a gene in both samples and comparing the		
XX	expression of the gene in both soft tissue samples, where a higher level		
XX	of protein expression in the first soft tissue sample indicates the		
XX	presence of soft tissue sarcoma. The method of the invention has		
XX	cytostatic applications and may be useful for detecting soft tissue		
XX	sarcoma, possibly via gene therapy or vaccine production. The nucleic		
XX	acid sequences may be useful in diagnostic and screening applications.		
XX	The current sequence is that of a human soft tissue sarcoma-upregulated		
XX	DNA of the invention. The current sequence is not shown within the		
XX	specification per se but was submitted in CD format by the inventor.		
XX	Sequence 518 BP; 116 A; 146 C; 149 G; 107 T; 0 U; 0 Other;		
SO			
Query Match	42.8%; Score 454; DB 12; Length 518;		
Best Local Similarity	92.3%; Pred. No. 1.5e-95;		
Matches	478; Conservative 0; Mismatches 40; Indels 0; Gaps 0		
QY	520 ACCCCTGCGCATCCGAGACAGTGCCTTCCCATGAGTCCATTGAGATTACCGTGAACGTT	579	
DB	518 ACTCCATGACATCTTCAATGCTCCGTCCTTCCCATTAAGTACCATGATTAATCTGACCGTT	459	
QY	580 CCGAGAGACGGGAGAGCGCAGAAAGCGTCTCTGATGACAGCCGGAGATGTGAATGTG	639	
DB	458 CCCAGAGCGGGCAGAGTGCATTAACGCTCTGTATGCGCCGCCCTAATGTGCAATTG	399	
QY	640 TCCGAGAACTGCATCTCTGAGCGGCTTAAGACTGAGCCTGCGCCTTGAATGCCAGAG	699	
DB	398 TCCAGGCAATGTGACTCTCTGAGAAAGGCTTAAGACTGAGCCTGCGCCTTGAATGCCAGAG	339	
QY	700 GCAGAGAACTGAGAGAAAGGCGGCTTCCATTAACGAAATCTGACAGAGCTGAATCTGA	759	
DB	338 GCAGAGTAAGTGAAGAGAAAGGCGGCTTCCATTAACGAAATCTGACAGAGCTGAATCTGA	279	
QY	760 GGGCTCTGAGAGCGCAGTCTGTCTCTGAAACAGAGCTTGCCTGGAGACGAGTGAAGTGGC	819	
DB	278 GGGCTCTGAGAGCGCAGTCTGTCTCTGAAACAGAGCTTGCCTGGAGACGAGTGAAGTGGC	219	
QY	820 AGCTGAAGTGGCTCTGGAGTCTTCAACAGGCGTTCGCTTGCCTTGAAGCTTGAACAA	879	

Db 218 AGCTGAAGTGGCTCTGGGGTCTCATGAGGCTGCTGCTCCAGCTGACACA 159
 Qy 880 GCCTGAGAAATCCCGCCGTAATATATCACTTTGGGGTTGGAGCTGTGTCCTCCGAC 939
 Db 158 GCGTGAAGAAATCCCGCCGTAATATATCACTTTGGGGTTGGAGCTGTGTCCTCCGAC 99
 Qy 940 GCTCTGACCTTCTGACGACGACCTGAGAAATGACCTGCTGCTGCTGCTGCTGCTG 999
 Db 98 GCTCTGACCTTCTGACGACGACCTGAGAAATGACCTGCTGCTGCTGCTGCTGCTG 39
 Qy 1000 TAAATGAATAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1037
 Db 38 TAAATGAATAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1

RESULT 6
 ADI31239
 ID ADI31239 standard; cDNA; 528 BP.

AC ADI31239;
 DT 17-JUN-2004 (first entry)
 XX
 DE Human cDNA #565.
 XX
 XX Human; gene; ss; immunological response; immunopathological condition;
 KM Crohn's disease; asthma; ulcerative colitis; hypersensitivity;
 KM irritable bowel syndrome; osteoarthritis; rheumatoid arthritis;
 KM acute monocytic leukemia; antineoplastic; antineoplastic; antineoplastic;
 KM osteoplastic; antineoplastic; antineoplastic; antineoplastic; antineoplastic;

OS Homo sapiens.
 XX US6607879-B1.
 XX 19-AUG-2003.
 XX 09-FEB-1998; 98US-00023655.
 XX 09-FEB-1998; 98US-00023655.
 PA (INCYTE) INCYTE CORP.
 PI Cocks BG, Stuart SG, Seilhamer JJ;
 DR WPI; 2003-895307/82.

PT A composition comprising a plurality of cDNAs, useful for detecting
 PT altered expression of genes in an immunological response or for
 PT diagnosing and treating an immunopathology, e.g. Crohn's disease, asthma
 PT or osteoarthritis.

PS Claim 1; SEQ ID NO 565; 50pp; English.

CC The invention relates to a composition comprising a plurality of cDNAs
 CC for detecting the altered expression of genes in an immunological
 CC response. The invention also relates to a method of diagnosing or
 CC monitoring the treatment of an immunopathological condition in a sample,
 CC comprising obtaining nucleic acids from a sample, contacting the nucleic
 CC acids of the sample with an array comprising the plurality of cDNAs under
 CC conditions to form one or more hybridisation complexes, detecting the
 CC hybridisation complexes and comparing the levels of the detected
 CC hybridisation complexes with the level of hybridisation complexes
 CC detected in a non-diseased sample, where an altered level of the detected
 CC hybridisation complexes correlates with the presence of an
 CC immunopathological condition. Also disclosed are an expression profile
 CC comprising a microarray and a plurality of detectable complexes and a
 CC method for identifying a plurality of polynucleotide probes. The cDNAs
 CC are useful as hybridisable array elements in a microarray for monitoring
 CC the expression of target polynucleotides. The microarray can be used in
 CC the diagnosis of an immunopathology, such as Crohn's disease, asthma,
 CC ulcerative colitis, hypersensitivity, irritable bowel syndrome,
 CC osteoarthritis, rheumatoid arthritis or acute monocytic leukemia, and in

CC identifying agents for the treatment of the diseases. The microarray may
 CC also be used in drug discovery and development, toxicological and
 CC carcinogenicity studies, forensics or pharmacogenomics. The composition
 CC may also be used in purification of a subpopulation of mRNAs, cDNAs or
 CC genomic fragments. This sequence represents a human cDNA of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification but was obtained in electronic format directly
 CC from USPTO at seqdata.uspto.gov/sequence.html.

SO Sequence 528 BP; 97 A; 138 C; 167 G; 126 T; 0 U; 0 Other;

Query Match 39.3%; Score 416.8; DB 11; Length 528;
 Best Local Similarity 99.3%; Pred. No. 6; e-87;
 Matches 429; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 599 CAGAGGCTCTGATGAGGAGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 658
 Db 1 CAGAGGCTCTGATGAGGAGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 60
 Qy 659 GAGCGGCTTAAGACTGAGCTGCGCCCTGAGTTCCAGAGGAGAGAGAGAGAGAGAGAGAGAG 718
 Db 61 GAGCGGCTTAAGACTGAGCTGCGCCCTGAGTTCCAGAGGAGAGAGAGAGAGAGAGAGAGAG 120
 Qy 719 AGGAGGCTCCAGATTAAGAGATTTGAGAGAGCTGAATCTAGAGGCTGTGAGAGCCAGT 778
 Db 121 AGGAGGCTCCAGATTAAGAGATTTGAGAGAGCTGAATCTAGAGGCTGTGAGAGCCAGT 180
 Qy 779 CTGCTCTGAGAACAGGCTTCTGAGG-ACGCTGAGCTGAGGAGAGAGAGAGAGAGAGAGAGAG 837
 Db 181 CTGCTCTGAGAACAGGCTTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
 Qy 838 GGTCTCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 897
 Db 241 GGTCTCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
 Qy 898 TAACTTATATCACTTTGGGGTTGCGCTGTGTCCTCCGAAACGCTGCACTTTGACG 957
 Db 301 TAACTTATATCACTTTGGGGTTGCGCTGTGTCCTCCGAAACGCTGCACTTTGACG 360
 Qy 958 CAGCTGAGAAATGACCTGCGCCCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1017
 Db 361 CAGCTGAGAAATGACCTGCGCCCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 Qy 1018 CGTGTGTCTGTG 1029
 Db 421 CGTGTGTCTGTG 432

RESULT 7
 ID AAX89074 standard; DNA; 1260 BP.

XX AAX89074;

DT 14-SEP-1999 (first entry)

DE Murine LAT (linker for activation of T cells) protein encoding DNA.

KM LAT; tyrosine kinase; linker for activation of T cell; TCR; human;
 KM T-cell receptor; TCR signaling pathway; neoplasia; inflammation;
 KM hypersensitivity; allergy; microbial infection; genetic disease;
 KM autoimmune disease; graft rejection; modulator; mouse; ss.

OS Mus musculus.

PN W09932627-A2.

PD 01-JUN-1999.

PF 23-DEC-1998; 98WO-US027400.

PR 23-DEC-1997; 97US-0068690P.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX SameLson LE, Zhang W;

XX WPI, 1999-418926/35.

DR P-PSDB; AAY27121.

PT Linker for activation of T cell protein used to, e.g. screen for
PT modulators of T cell signaling.

PS Example 2, Fig 7C, 125pp; English.

XX The invention relates to a protein tyrosine kinase substrate LAT (linker
XX for activation of T cells) protein. Modulation of interaction between LAT
XX and the T-cell receptor (TCR) affects the TCR signaling pathway. LAT is
XX a substrate for tyrosine kinases and becomes phosphorylated after TCR
XX engagement, resulting in recruitment of other signaling molecules. LAT
XX is used to identify and test (ant)agonists of tyrosine kinase signaling
XX pathways, i.e. modulation of interaction between tyrosine kinase
XX substrates and intracellular ligands or between these ligands and other
XX members of the pathway, including identification of downstream signalling
XX proteins, particularly in immune system cells. These modulators are
XX potentially useful as drugs and diagnostic agents, particularly for
XX diseases that involve undesirable cell proliferation, differentiation,
XX growth or T cell anergy, e.g. neoplasia, inflammation, hypersensitivity/
XX allergy, microbial infection, metabolic, genetic or autoimmune diseases,
XX graft rejection. LAT is also used to generate specific antibodies, used
XX for detection of LAT. Nucleic acid that encodes LAT, or its fragments,
XX are used to identify homologous sequences in other species; to detect the
XX LAT gene and as sources of antisense therapeutics. Modulators of LAT are
XX potentially more specific and less toxic than known immunosuppressants
XX such as cyclosporin. The present sequence represents the nucleotide
XX sequence of murine LAT

SQ Sequence 1260 BP; 282 A; 359 C; 360 G; 259 T; 0 U; 0 Other;

Query Match 37.8%; Score 400.4; DB 2; Length 1260;

Best Local Similarity 67.6%; Pred. No. 5,5e-83;

Matches 698; Conservative 0; Mismatches 291; Indels 43; Gaps 8;

QY 26 GTGACGACGAGCTGCTCCGAGCTCCCTGAGATGAGAGGAGCCATCTGTCCTCGG 85
DB 260 GTGACGACGAGGAGCTGAGAGCTGCTGACATGAGAGAGAGGAGCCGCTGAGAGGAGT 319
QY 86 TGGTGGGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 142
DB 320 GGTGGGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 379
QY 143 GGCACAGATGCGAGGCTCTGAGACAGACATCCATGTTGATCCAGAGGAGGCA 202
DB 380 GCGGTGAGTGGGAGTCTCTCTGAGACAGACATCCAGAGAGGTTGATCCAGAGAGCA 439
QY 203 TCCAGTTCAAGGAGGCTGACAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 256
DB 440 TCCATCATGAGGAGCTGATTAACCGTCCCGAAGACATCTGCTGCTGCTGCTGCTGCTGCTGCT 499
QY 257 TCACTCTCTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 316
DB 500 TCACT 559
QY 317 CCTTGGGGGGGCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 376
DB 560 CCTTGGGGGGGCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 619
QY 377 TGGGAGGAGTACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 436
DB 620 TGGGAGGAGTACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 679
QY 437 ATCAACAGCC--AGGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 493
DB 680 AGGACTATCTCCAGAGGAGTACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 739
QY 494 CTGCCCATCAGCTCTGCACTCAGAGCCCTGGAGATCCGAGAGAGTGCTCTCTCATAGG 553

DB 740 TTGTCT 799
QY 554 AGTCCATTGATGATTAACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 813
DB 800 AGTCTGTGTAATTAACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 859
QY 614 ATGGAG 673
DB 860 ATGGAG 916
QY 674 AGCTGCGGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 733
DB 917 AGCTGCGGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 976
QY 734 AGGAGATCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 793
DB 977 GAGAG 1036
QY 794 GGTGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 853
DB 1037 TGTGTCCTG-----AACTGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1088
QY 854 CTGCGCTTGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 913
DB 1089 CTGCGCTTGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1136
QY 914 TGGGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 973
DB 1137 TGGGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1194
QY 974 TGGGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1033
DB 1195 CTGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1248
QY 1034 AAAAAAAAAA 1045
DB 1249 AAAAAAAAAA 1260

RESULT 8
AAK57405
ID AAK57405 strand1; cDNA; 312 BP.
XX
AC AAK57405;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:2465.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ss.
OS Homo sapiens.
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-018464P.
PR 16-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214865P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216860P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241836P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249255P.
PR 17-NOV-2000; 2000US-0249257P.
PR 17-NOV-2000; 2000US-0249259P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250393P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251031P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI: 2001-463426/52.
P-PSDB; AAM64624.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
useful for preventing, diagnosing and/or treating cancers and metastasis.

Claim 1; SEQ ID NO 2465; 3071bp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
activity, and can be used in gene therapy and vaccine production. (I)
proteins and polynucleotides may be used in the prevention, diagnosis and
treatment of diseases associated with inappropriate (I) expression. For
example, they may be used to treat disorders associated with decreased
expression by rectifying mutations or deletions in a patient's genome
that affect the activity of (I) by expressing inactive proteins or to
supplement the patients own production of (I). Additionally, (I)

CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 312 BP; 64 A; 113 C; 70 G; 64 T; 0 U; 1 Other;
Query Match 7.8%; Score 83; DB 4; Length 312;
Best Local Similarity 89.3%; Pred. No. 1.8e-09;
Matches 134; Conservative 0; Mismatches 11; Indels 5; Gaps 4;
QY 303 AAGATCCCGCGGACCCCTTTGGGGGCTCCACCGGAGCCATTTCCCGCGGGA-TTCTG 361
DB 163 AAGATCCCGCGGACCCCTTTGGGGGCTCCACCGGAGCCATTTCCCGCGGATTCTG 222
QY 362 ATGATGCCAAGCTGTGGGAGCT-ACGAGAAGAGAAGCAAGCCTGTG-AGATGACAGA 419
DB 223 ATGATGCCAAGCTGTGGGAGCTAAGAGAAGAGAAGCAAGCCTGTGAGAGATGCAGA 282
QY 420 TGA--GGATGAGAGAGACTATCAACACCA 447
DB 283 TGAAGATGAGAGAGAGACTATCAACANCA 312
RESULT 9
AAK57403
ID AAK57403 standard; cDNA; 519 BP.
XX AAK57403;
AC
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:2463.
XX
KM Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX
OS cytostatic; gene therapy; vaccine; metastasis; ss.
XX
XX Homo sapiens.
XX
XX MO200157182-A2.
XX
XX
PD 09-AUG-2001.
XX
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216890P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227099P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.

PR	08-NOV-2000	2000US-0246527P.
PR	08-NOV-2000	2000US-0246528P.
PR	08-NOV-2000	2000US-0246532P.
PR	08-NOV-2000	2000US-0246509P.
PR	08-NOV-2000	2000US-0246610P.
PR	08-NOV-2000	2000US-0246611P.
PR	08-NOV-2000	2000US-0246613P.
PR	17-NOV-2000	2000US-0249207P.
PR	17-NOV-2000	2000US-0249208P.
PR	17-NOV-2000	2000US-0249209P.
PR	17-NOV-2000	2000US-0249210P.
PR	17-NOV-2000	2000US-0249211P.
PR	17-NOV-2000	2000US-0249212P.
PR	17-NOV-2000	2000US-0249213P.
PR	17-NOV-2000	2000US-0249214P.
PR	17-NOV-2000	2000US-0249215P.
PR	17-NOV-2000	2000US-0249216P.
PR	17-NOV-2000	2000US-0249217P.
PR	17-NOV-2000	2000US-0249218P.
PR	17-NOV-2000	2000US-0249244P.
PR	17-NOV-2000	2000US-0249245P.
PR	17-NOV-2000	2000US-0249264P.
PR	17-NOV-2000	2000US-0249265P.
PR	17-NOV-2000	2000US-0249297P.
PR	17-NOV-2000	2000US-0249299P.
PR	17-NOV-2000	2000US-0249300P.
PR	01-DEC-2000	2000US-0250160P.
PR	01-DEC-2000	2000US-0250391P.
PR	05-DEC-2000	2000US-0251030P.
PR	05-DEC-2000	2000US-0251988P.
PR	05-DEC-2000	2000US-0256719P.
PR	06-DEC-2000	2000US-0251479P.
PR	08-DEC-2000	2000US-0251856P.
PR	08-DEC-2000	2000US-0251868P.
PR	08-DEC-2000	2000US-0251869P.
PR	08-DEC-2000	2000US-0251889P.
PR	08-DEC-2000	2000US-0251990P.
PR	11-DEC-2000	2000US-0254097P.
PR	05-JAN-2001	2001US-0259678P.

PA (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI: 2001-483426/52.
DR P-PSDB; AAM84622.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis
XX
PS Claim 1, SEQ ID NO 2463, 3071pp + Sequence Listing; English.

CC AA65495 to AA664702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AA654703
CC to AA687694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AA654942 to AA654950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX

SQ Sequence 519 BP; 80 A; 156 C; 161 G; 118 T; 0 U; 4 Other;

Query Match 6.4%, Score 68; DB 4; Length 519;

[illegible]

RESULT 10
 ADAL19460
 ID ADAL19460 standard; DNA; 6307 BP.
 XX
 AC ADAL19460;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Mouse LAT Y136F gene SEQ ID NO:1.
 KW mutant; linker for activation of T-cells; LAT; TH2 cell differentiation;
 KW allergy; asthma; eosinophilia; TH2 cell deregulation; antiasthmatic;
 KW antiallergic; gene; ds.
 XX
 OS Mus musculus.
 XX
 PN WO2003068968-A2.
 XX
 PD 21-AUG-2003.
 XX
 PF 14-FEB-2003; 2003WO-IB001044.
 XX
 PR 14-FEB-2002; 2002US-0356136P.
 PR 11-MAR-2002; 2002EP-00290610.
 XX
 XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 PA
 XX
 PI Malissen M, Malissen B, Aguado Vidal E;
 XX
 DR MPI; 2003-679639/64.
 XX
 PT New non-human animal (e.g. mouse) comprising a mutated linker for
 PT Activation of T-cells (LAT) gene coding for a mutant LAT protein, useful
 PT as a model of allergy, asthma, eosinophilia or a disease associated with
 PT TH2 cell deregulation.
 XX
 PS Claim 30; Page 46-50; 52pp; English.
 CC
 CC The present invention describes a non-human animal having a mutated
 CC linker for activation of T-cells (LAT) gene coding for a mutant LAT
 CC protein, where the mutant LAT protein leads to an exaggerated TH2 cell
 CC differentiation. Also described: (1) a germ cell or somatic cell from a
 CC non-human animal or its progeny containing the mutated LAT gene; (2)
 CC screening for a drug for creating allergy, asthma and/or disease
 CC associated with TH2 cell deregulation, or for drugs that regulate the
 CC activity of TH2 cells; (3) producing a pharmaceutical composition for
 CC treating a disease associated with deregulated TH2 cells activity; (4) a
 CC bioreactor for a large scale production of human IGE antibodies
 CC comprising an animal; (5) producing human IGE antibodies; (6) a B-cell
 CC hybridoma obtained by the above method; (6) a mutated mouse gene coding
 CC for a mutant LAT protein, the sequence of which corresponds to a wild
 CC type sequence and contains a single mutation of the tyrosine Y136 or a
 CC composite mutation of the three distal tyrosine residues; (7) a
 CC diagnostic method for asthma, allergy, eosinophilia and/or TH2 cells
 CC deregulation; (8) a diagnostic kit for asthma, allergy, eosinophilia
 CC and/or TH2 cells deregulation, comprising oligonucleotide probes for the
 CC detection of a mutated LAT gene coding for a mutant LAT protein
 CC containing a single mutation of the tyrosine Y136 or a composite mutation
 CC of the three distal tyrosine residues; and (9) a non-human animal
 CC resulting from the breeding of a non-human animal expressing humanized
 CC IGE with the non-human animal described above. Mutant LAT has

CC antiasthmatic and antiallergic activities. The animal is useful as a
CC model of allergy, asthma, eosinophilia and/or a disease associated with
CC TH2 cell deregulation. The methods may be used for screening drugs that
CC may be utilized for treating asthma or allergy or for diagnosing the
CC above diseases. The present sequence represents a mutated mouse LAT gene
CC which encodes a LAT Y136F protein.

XX Sequence 6307 BP; 1219 A; 1676 C; 1556 G; 1856 T; 0 U; 0 Other;

Query Match 5.2%; Score 54.6; DB 9; Length 6307;

Best Local Similarity 71.3%; Pred. No. 0.017;

Matches 72; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Oy 457 GTGGTCTCTCTGACAGACACCCGCGCACTAGACATCTGCCCCCATCTCTGCACTC 516

Db 2048 GTGGTCTCTCTGACAGATGCTCTGCTGCGCTCTGCTCTCTCTCTCTGCTCT 2107

Oy 517 AGCACCCCTGGCATCCGACAGATGCTCTCTGATGAGTCT 557

Db 2108 AGCAACCTGACCTTGAGACAGTGTCTCTGAGTCTC 2148

RESULT 11

ADA71938

ID ADA71938 standard; DNA; 2000 BP.

XX ADA71938;

DT 20-NOV-2003 (first entry)

DE Rice gene, SEQ ID 5263.

KW Plant; bacterial infection; fungal infection; viral infection; rice;

KM gene; db.

OS Oryza sativa.

PN MO2003000898-A1.

PD 03-JAN-2003.

PF 22-JUN-2001; 2001WO-1B001105.

PR 22-JUN-2001; 2001WO-1B001105.

PA (SYGN) SYNGENTA PARTICIPATIONS AG.

PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

DR WPI; 2003-175290/17.

PT Identifying at least one gene involved in plant resistance or response to

PT pathogenic infection for conferring resistance or tolerance to a plant to

PT bacterial, fungal or viral infection by determining or detecting plant

PT gene expression.

XX Claim 27; SEQ ID NO 5263; 8999P; English.

XX The present invention relates to a method (M1) for identifying genes

XX involved in plant resistance or response to pathogenic infection. M1

XX comprises identifying a gene whose expression is significantly altered in

XX the incompatible interaction of plant gene expression relative to

XX expression of the gene in an uninfected plant, in a mutant plant that

XX does not express a gene associated with response to pathogenic infection,

XX or in a corresponding incompatible or compatible interaction. (M1) is

XX useful for conferring resistance to resistance or tolerance to a plant to

XX bacterial, fungal or viral infection. The present sequence was used to

XX illustrate the invention.

XX Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

XX Query Match 5.0%; Score 53.2; DB 8; Length 2000;

Best Local Similarity 10.1%; Pred. No. 0.026;

Matches 76; Conservative 355; Mismatches 308; Indels 11; Gaps 4;

Oy 313 CAGCCCTTGGGGGCTCCACCGACGCGCATTTCCCGGGGATTTGATGTGCAAC 372

Db 23 MAGMMSCARMGSSSRKMSKRSKRSKRSKRSKRSKRSKRSKRSKRSKRSKRS 82

Oy 373 AGTGTGGAGCTAG 432

Db 83 GSYSGKMKKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYK 142

Oy 433 GACTATCA-----CAACCGAGCTACTGCTGCTCTCTTCAACAGACACCCGCGCACTA 487

Db 143 SYMMVCYARCGSCCRKSKSKSGSWKTCRARGSGMSGSKSKSKSKSKSKSKSK 202

Oy 488 GCATCTGCGCCCATCAGCTCTGCACTAG--CAACCTGCACTCCAGACAGTGCCTT 545

Db 203 SCGGRSASVSRYYGTSRKYGTYSASCKMAVMTTSYSMASSSYWCRSKRSKM 262

Oy 546 CTCGATGAGCTCATTTAGTATGATTAAGCTTCGAGAGAGGAGAGAGAGAGAG 605

Db 263 WDMRKMRSRSGYMSYKMMCTAYKSKSYSRWCTYRGGMGATRYWGRGYSKM 322

Oy 606 GTCTCTGATGCGACCGGAGATGATGATGATGATGATGATGATGATGATGATG 665

Db 323 AMYKMMYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYK 382

Oy 666 TAAAGCTAGAGCTGCGCCCTGAGTTCCAGAGAGAGAGAGAGAGAGAGAGAGAG 725

Db 383 SRKKRKMCRKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRY 442

Oy 726 TCCAGATTACGAGATCTGACGAGACCTAAGCTGAGGCTGTGAGGCGCATCTGTCT 785

Db 443 RWSKRMASKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRY 502

Oy 786 GGAACGAGCTTGCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 845

Db 503 RRMGMRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRY 562

Oy 846 CATGCGCTC--CTGCCCTTGTCTGAGCTGAGAGAGAGAGAGAGAGAGAGAG 902

Db 563 YMYKMMYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRY 622

Oy 903 TATTTACTTGGGCTTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 962

Db 623 CASKRSARRYAMWMTSGSRMSKRYKRYKRYKRYKRYKRYKRYKRYKRYKRY 682

Oy 963 TG-AGAAAGAGCTGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1021

Db 683 WRIRYRAMCMYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRY 742

Oy 1022 TGTCTGTGAG 1051

Db 743 YRSMYVYMSWMAKTKMRYATRMWMMW 772

RESULT 12

AACT8190

ID AACT8190 standard; cDNA; 2968 BP.

XX AACT8190;

DT 08-FEB-2001 (first entry)

DE Human cancer associated gene sequence SEQ ID NO:584.

XX Human cancer associated gene; cancer antigen; detection; cancer;

XX diagnosis; cytostatic; proliferative; vulnery; immunomodulator;

XX antidiabetic; antiaesthatic; antineumatic; antiallergic; antiviral;

XX antineumatic; antineumatic; antineumatic; antineumatic; antineumatic;

XX dermatological; neoplastic; thrombolytic; coagulant; neoplastic;

XX vasotrophic; antineumatic; antineumatic; antineumatic; antineumatic;

XX immune disorder; haematopoietic cell disorder; autoimmune disorder;

KW Allergic reaction; graft versus host disease; organ rejection;
 KM haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO20005350-A1.
 PN
 XX
 PD 21-SEP-2000.
 PF
 PF 08-MAR-2000; 2000WO-US005882.
 XX
 XX 12-MAR-1999; 99US-0124270P.
 PR
 XX (HUMA-) HUMAN GENOME SCT INC.
 PA
 XX
 XX Rosen CA, Ruben SM;
 PI
 XX
 DR WPI: 2000-587533/55.
 DR P-PSDB; AAB43981.
 XX
 PT Novel isolated nucleic acids comprising sequences encoding peptides
 PT useful for treating or diagnosing e.g. cancer.
 XX
 PS Claim 1; Page 1106-1107; 2352pp; English.

CC AAB77607 to AAB78448 encode the human cancer associated proteins given in
CC AAB77607 to AAB78448. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vlnetary; immunomodulator;
CC antidiabetic; antiaesthetic; antineuritic; antiarteritic;
CC antihistaminic; antihypertensive; antiallergic; antibacterial; antiviral;
CC dermatological; neuroprotective; cardiac; thrombolytic; coagulant;
CC neurotropic; vasotropic; antipsoriatic and antineoplastic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haemotopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAB78449 to
CC AAB78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention

Sequence 2968 BP; 678 A; 817 C; 797 G; 671 T; 0 U; 5 Other;

Query Match 4.7%; Score 49.6; DB 3; Length 2568;
 Best Local Similarly 53.6%; Pred. No. 0.2;
 Matches 103; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY	869	GCCTGACAAAGCCCTGAGAAATATCCCCCGTACTTATATACCTTGGGGTTCCGCCCTGT	928
Db	2733	GCCCGGAGACAGCCCTCTTCGCAATGCGACATGATTTTGGCATTTACTTTAAAGCTCACCTTTT	2792
QY	929	GTCCCCCGGAAAGCTCTGACCTTCTGACGACGACCTGAGATATGACCTGCCCTGGCCCCAGC	988
Db	2793	TTCTTTCCCTCTGCTGTTGCTGCTGCTGACGATTAATGATTGTTGTTCTTCCCATGGGATC	2852
QY	989	CCTACTCTGTGTATATGATATTAAGGCTGCGCTGTGTCTGTGGAAAAAAAAAAAAAAAAAAAA	1048
Db	2853	CATCTGTTTGTAAACAATATAAGCGCTGAGAGGAGTGTAAAAAAAAAAAAAAAAAAAAAA	2912
QY	1049	AAAAAAAAAAAAA	1060
Db	2913	AAAAAAAAAAAAA	2924

RESULT 13
AAH34408

ID	AA34408	standard; cDNA; 2968 BP.
XX		
AC	AA34408;	
XX		
DT	03-SEP-2001	(first entry)
XX		
DE	Human colon cancer antigen encoding cDNA SEQ ID NO:1490.	
XX		
KW	Human; colon cancer; colon cancer antigen; diagnosis; detection;	
KW	colorectal carcinoma; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200122920-A2.	
XX		
PD	05-APR-2001.	
XX		
PP	28-SEP-2000; 2000WO-US026524.	
XX		
PR	29-SEP-1999; 99US-0157137P.	
PR	03-NOV-1999; 99US-0163280P.	
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Ruben SM, Barash SC, Birse CE, Rosen CA;	
XX		
DR	WPI; 2001-235357/24.	
DR	P-PSDB; AAG75003.	
XX		
PT	Nucleic acids encoding 4277 human colon cancer-associated polypeptides,	
PT	useful for preventing, diagnosing and/or treating colorectal cancers.	
XX		
SS	Claim 1; Page 3154-3155; 9803JP; English.	

AAH32943 to AAH71715 and AAC73514 to AAC77788 represent human colon cancer-associated nucleic acid molecules (N) and proteins (P), where the proteins are collectively known as colon cancer antigens. The colon cancer antigens have cytostatic activity and can be used in gene therapy and vaccine production. N and P may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate P expression. For example, N and P may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of P by expressing inactive proteins or to supplement the patient's own production of P. Additionally, N may be used to produce the colon cancer-associated Ps, by inserting the nucleic acids into a host cell and culturing the cell to express the proteins. N and P can be used in the prevention, diagnosis and treatment of colorectal carcinomas and cancers. AAH7196 to AAH37204 and AAC7789 represent sequences used in the exemplification of the present invention. N.B. Pages 666 to 682 and page 7053 of the sequence listing were missing at time of publication, meaning no sequences are present for SEQ ID NO:1027 to 1052, 7921 and 7922

SQ Sequence 2968 BP; 678 A; 817 C; 798 G; 671 T; 0 U; 4 Other;

Query Match	4.7%	Score	49.6	DB	4	Length	2968
Best Local Similarity	53.6%	Pred. No.	0.2				
Matches	103	Conservative	0	Mismatches	89	Indels	0
						Gaps	0

Oy	869	GCCGACAAAGCCTGTGAAGAATCCCCCGTAACATTATATCACTTTGGGGTTCGGCCTGT	928
Db	2733	GCCCGAAGACGCCTCTTGCGATAGCAGATGATTGGCATTTACTTTAAAGCTCACCTTTT	2792
Oy	929	GTECCCGGAACGCTCTGCACCTTCTTGACGACGCTGTGAATGACCTCCCTGGCCCCAGC	988
Db	2793	TTCCTCCCCTCTCGTGGCTGCTGTGAGATAATGATTGGTTCCTTCCCTAATGGATC	2852
Oy	989	CCTACTCTGTGTAAATAGATTAAGAAGCCTGCGCTGTGTCTGTGAAAAAAAAAAAAAAAAA	1048
Db	2853	CATCTGTTTTGTAAACAATAAAGCTGTGAGGAGAGTCTAAAAAABAAAAAAAAAAAAAAAA	2912
Oy	1049	AAAAAAAAAAAAA 1060	

CC	directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 375 BP; 130 A; 87 C; 65 G; 65 T; 0 U; 28 Other;

Query Match	4.6%	Score	48.6	DB	4	Length	375
Best Local Similarity	60.0%	Pred. No.	0.19				
Matches	78	Conservative	0	Mismatches	52	Indels	0
						Gaps	0

4.68; Score 48.6; DB 4; Length 375;
CO 08; Prod NO 018.

```

best local similarity 80.0%;  E-vec. no. 0.15;
Matches 78; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

```

```

best local similarity 80.0%;  E-vec. no. 0.15;
Matches 78; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

```

931 CCCCCGACGCTCTGCACCTTCTGACGCGAGCCTGAGAATGACCTGCCCTGGCCCCAGCCC 990

Db 197 CCCCTGTAATCCCGAGCACTTGGGAGGCTGCCGAGAACTGCTTGGACCCAGTGCAAAAC 256

QY 991 TACTCTGTGTAATAGATAAAGGCTGCGTGTGTCTGTGAAAAA 1050

Db 257 CATTCTGGCACAATAGCGAGACCATGTCTGTNNANAAAAAAAAA 316

QY 1051 AAAAAAAAAA 1060
|||||||

Db 317 AAAAAAAAAA 326

Search completed: November 22, 2004, 08:24:08
Job time : 889 secs

Job time : 889 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 07:49:15 ; Search time 208 Seconds
(without alignments)
3622.289 Million cell updates/sec

Title: US-09-597-920B-1

Perfect score: 1060

Sequence: 1 gactctgccttcgagggcc.....aaaaaaaaaaaaaaaa 1060

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents, NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	416.8	39.3	528	US-09-023-655-565	Sequence 565, App
2	51.4	4.8	7218	US-08-232-463-14	Sequence 14, Appl
3	43.8	4.1	1926	US-09-249-585A-4	Sequence 4, Appl
4	43.8	4.1	1931	US-09-130-114-2	Sequence 2, Appl
5	42.4	4.0	1503	US-09-907-794A-220	Sequence 220, App
6	42.4	4.0	1503	US-09-905-125A-220	Sequence 220, App
7	42.4	4.0	1503	US-09-902-775A-220	Sequence 220, App
8	42.4	4.0	1503	US-09-906-700-220	Sequence 220, App
9	42.4	4.0	1503	US-10-140-002-363	Sequence 363, App
10	42.4	4.0	1503	US-09-903-603A-220	Sequence 220, App
11	42	4.0	977	US-09-311-021-103	Sequence 103, App
12	42	4.0	3083	US-09-155-246-1	Sequence 1, Appl
13	42	4.0	4494	US-09-620-312D-861	Sequence 861, Appl
14	41.8	3.9	1505	US-07-915-246-1	Sequence 1, Appl
15	41.8	3.9	1600	US-07-861-458C-37	Sequence 37, Appl
16	41.8	3.9	1607	5196333-3	Patent No. 5196333
17	41.6	3.9	1160	US-08-995-159-1	Sequence 1, Appl
18	41.6	3.9	1160	US-09-545-605-1	Sequence 1, Appl
19	41.2	3.9	966	US-08-921-382-1	Sequence 1, Appl
20	41.2	3.9	966	US-09-386-380-1	Sequence 1, Appl
21	41.2	3.9	3489	US-08-728-323A-1	Sequence 1, Appl
22	41.2	3.9	3489	US-09-298-568-1	Sequence 1, Appl
23	41.2	3.9	3489	US-09-410-399-1	Sequence 1, Appl
24	41.2	3.9	3489	US-09-894-273-1	Sequence 1, Appl
25	41.2	3.9	32207	US-08-770-379-20	Sequence 20, Appl
26	41.2	3.9	32207	US-08-757-669A-20	Sequence 20, Appl
27	41.2	3.9	32207	US-09-230-371A-20	Sequence 20, Appl

28	41	3.9	1366	US-09-200-965-1	Sequence 1, Appl
29	40.8	3.8	3028	US-09-548-938A-2	Sequence 2, Appl
30	40.4	3.8	997	US-09-800-729-14	Sequence 14, Appl
31	40.4	3.8	1071	US-09-800-729-45	Sequence 45, Appl
32	40.4	3.8	1086	US-09-800-729-48	Sequence 48, Appl
33	40.4	3.8	1138	US-09-800-729-44	Sequence 44, Appl
34	40.4	3.8	1149	US-09-800-729-47	Sequence 47, Appl
35	40.4	3.8	1798	US-09-797-906-1	Sequence 1, Appl
36	40.4	3.8	7218	US-08-232-463-14	Sequence 14, Appl
37	40.4	3.8	16442	US-08-781-891-208	Sequence 208, App
38	40.4	3.8	16442	US-09-618-166-208	Sequence 208, App
39	40.2	3.8	1851	US-09-149-476-62	Sequence 62, Appl
40	40.2	3.8	2920	US-09-276-400-1	Sequence 1, Appl
41	40.2	3.8	2920	US-09-448-076-1	Sequence 1, Appl
42	40.2	3.8	2920	US-09-702-572-1	Sequence 1, Appl
43	40	3.8	2744	US-09-071-101-1	Sequence 1, Appl
44	40	3.8	2744	US-09-369-618-1	Sequence 1, Appl
45	40	3.8	2744	US-09-369-617-1	Sequence 1, Appl

ALIGNMENTS

```
RESULT 1
US-09-023-655-565
; Sequence 565, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 565:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 528 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: HIPNOTO1
; CLONE: 240885
; US-09-023-655-565
Query Match 39.3%; Score 416.8; DB 4; Length 528;
```


Best Local Matches	Similarity	99.3%	Pred. No. 8_9e-95	Conservative	0	Mismatches	2	Indels	1	Gaps	1
Oy	599	CAGAAAGCCTCTCTGAGATGCGACCCGGAGTATGTGATGTGTC	CCAGAACTGCATCCTG	658							
Db	1	CAGAAAGCCTCTCTGAGATGCGACCCGGAGTATGTGATGTGTC	CCAGAACTGCATCCTG	60							
Oy	659	GAGCGGCTAAGCTAGACCTGCGCGCCCTGAGATGTC	CCAGAAAGGCAAGAACTGAGAGAA	718							
Db	61	GAGCGGCTAAGCTAGACCTGCGCGCCCTGAGATGTC	CCAGAAAGGCAAGAACTGAGAGAA	120							
Oy	719	AGGGGCTCCAGATTACGAGAACTTGCAGAGACTGAATCGAGG	CGCTGAGAGCCGAGT	778							
Db	121	AGGGGCTCCAGATTACGAGAACTTGCAGAGACTGAATCGAGG	CGCTGAGAGCCGAGT	180							
Oy	779	CTGTCTTGGAACCAAGGCTTGCCTGGG-ACGGCTAGACTGGG	CAGCTGGAATGAGCTCTGG	837							
Db	181	CTGTCTTGGAACCAAGGCTTGCCTGGGAAACGGCTAGACTGGG	CAGCTGGAATGAGGCTCTGG	240							
Oy	838	GGTCTTCACATGGCGCTCTGCGCCCTTGTCTCCACCTGACA	CAACAGCTGAGAAATCCCCCG	897							
Db	241	GGTCTTCACATGGCGCTCTGCGCCCTTGTCTCCACCTGACA	CAACAGCTGAGAAATCCCCCG	300							
Oy	898	TAACTTATATCACTTTGGGGTTCGAGCTGTGTCCCGGAACG	CTGTGCACTTCTGAC	957							
Db	301	TAACTTATATCACTTTGGGGTTCGAGCTGTGTCCCGGAACG	CTGTGCACTTCTGAC	360							
Oy	958	CAGCCTGAGAAATGACCTGCGCTGAGCCCAAGCCCTA	CTGTGTAAATAGAAATAAAGCCTG	1017							
Db	361	CAGCCTGAGAAATGACCTGCGCTGAGCCCAAGCCCTA	CTGTGTAAATAGAAATAAAGCCTG	420							
Oy	1018	CGTGTGTCTGTG	1029								
Db	421	CGTGTGTCTGTG	432								

RESULT 2
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, P. G.
TITLE OF INVENTION: RECOMBINANT FOMLPX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION/DOCKET NUMBER: 29,768
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300

```

; TELEFAX: (703)683-4109
;
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 7218 base pairs
;
; TYPE: nucleic acid
;
; STRANDEDNESS: single
;
; TOPOLOGY: linear
;
; IMMEDIATE SOURCE:
;
; CLONE: pTZgpc-Fl6
;
US-08-232-463-14

```

[illegible]

```

RESULT 3
US-09-249-585A-4/c
; Sequence 4, Application US/09249585A
; Patent No. 6417002
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert
; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
; FILE REFERENCE: 0867/0D905
; CURRENT APPLICATION NUMBER: US/09/249,585A
; CURRENT FILING DATE: 1999-02-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein Barr Virus
; FEATURES:
; NAME/KEY: misc feature
; LOCATION: (1)..(1926)
; OTHER INFORMATION: template strand of EBNA-1 DNA
US-09-249-585A-4

```

	Query Match	Similarity	4.1%;	Score 43.8;	DB 4;	Length 1926;
	Best Local	Similarity	47.0%;	Pred. No. 0.2;	Mismatches 152;	Indels 0;
	Matches	135;	Conservative	0;	Gaps	0;
Oy	552	GGAGTCATTGATATTACGTGAACGTTCCGAGAGCGGGAGAGCGCCAGACGTTCTT	611			
Db	575	CGAGGACGGGAGAGAGACGACGACGGGGAGAGACGGGAGAGACGAGACGAGACGGGAGAGA	516			
Oy	612	GGATGGACGCGGGAGATGTGAATGTGTCCAGGAACCTGCATCTTGGACGGGCTAAGAC	671			

Db 515 GGACGAGGACGGGAGGACGGGAGGACGGGAGGACGAGACGGGAGGACGGGAG 456
Qy 672 TGAGCTGCGCGCCCTGATGTTCCAGGAGGACGAGAAAGTGAAGAGGGGCTCCAGA 731
Db 455 GGAGGACGAGGACGGGAGGAGGAGGACGAGGACGGGAGGACGGGAGGAGGAGGA 396
Qy 732 TTACGAGATCTGCAGAGCTGAAGCTGAGGGCTGTGAGGCGAGTCTGTCTTGAAAC 791
Db 395 CGAGGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 336
Qy 792 AGGCTTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 838
Db 335 GGACGAGGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 289

RESULT 4

US-09-130-114-2/c
; Sequence 2, Application US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Damaj, Bassem B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
; FILE REFERENCE: 0867/ID903US1
; CURRENT APPLICATION NUMBER: US/09/130,114
; CURRENT FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: EBNA
US-09-130-114-2

Query Match 4.1%; Score 43.8; DB 2; Length 1931;
Best Local Similarity 47.0%; Pred. No. 0.2;
Matches 135; Conservative 0; Mismatches 152; Indels 0; Gaps 0;

Qy 552 GGAGTCATTTGATGATTTACGTGAACGTTCCGAGAGCGGGAGAGCGGAGAAAGCTTCT 611
Db 575 GGAGGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 516
Qy 612 GGATGCGACCGCGGAGTATGTGAATGTCTCCAGGAACCTGCTCTGAGCGGCTTAAGC 671
Db 515 GGACGAGGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 456
Qy 672 TGAGCTGCGCGCCCTGATGTTCCAGGAGGACGAGAAAGTGAAGAGGGGCTCCAGA 731
Db 455 GGAGGACGAGGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 396
Qy 732 TTACGAGATCTGCAGAGCTGAAGCTGAGGGCTGTGAGGCGAGTCTGTCTTGAAAC 791
Db 395 CGAGGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 336
Qy 792 AGGCTTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 838
Db 335 GGACGAGGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 289

RESULT 5

US-09-907-794A-220
; Sequence 220, Application US/09907794A
; Patent No. 6635468
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Baton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen

APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/907,794A
CURRENT FILING DATE: 2001-07-17
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 220
LENGTH: 1503
TYPE: DNA
ORGANISM: Homo sapiens
US-09-907-794A-220

Query Match 4.0%; Score 42.4; DB 4; Length 1503;
Best Local Similarity 66.3%; Pred. No. 0.41;
Matches 61; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 969 TGACCTGCGCGCCCGCCAGCCCTACTCTGTGTAATGAATGAAGCGCTGCTGTCTGT 1028
Db 1385 TGGCTGCGCTGGGACACCTCTCTCCAGGAGCAATGAAGCGCGCGGAGACT 1444
Qy 1029 GGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1060


```

; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 220
; LENGTH: 1503
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-775A-220

Query Match
Best Local Similarity 66.3%; Pred. No. 0.41; Length 1503;
Matches 61; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 969 TGACCTGCCCTGGCCCGCCCTACTCTGTGTAATGAAATAAGCCCTGGCTGTCTGT 1028
Db 1385 TGGCGTGGCTGGGACACCTCTCTGCGACGAGGCATTAAGCCGCGGAGCCT 1444

Qy 1029 GGAATAAAAAAAAAAAAAAAAAAAAAA 1060
Db 1445 TGAATAAAAAAAAAAAAAAAAAAAAAA 1476

RESULT 8
US-09-906-700-220
; Sequence 220, Application US/09906700
; Patent No. 6723535
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan J.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
```

```

; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 220
; LENGTH: 1503
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-906-700-220

Query Match
Best Local Similarity 66.3%; Pred. No. 0.41; Length 1503;
Matches 61; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 969 TGACCTGCCCTGGCCCGCCCTACTCTGTGTAATGAAATAAGCCCTGGCTGTCTGT 1028
Db 1385 TGGCGTGGCTGGGACACCTCTCTGCGACGAGGCATTAAGCCGCGGAGCCT 1444

Qy 1029 GGAATAAAAAAAAAAAAAAAAAAAAAA 1060
Db 1445 TGAATAAAAAAAAAAAAAAAAAAAAAA 1476

RESULT 9
US-10-140-002-363
; Sequence 363, Application US/10140002
; Patent No. 6725730
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
```

APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P33081C59
CURRENT APPLICATION NUMBER: US/10/140,002
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 363
LENGTH: 1503
TYPE: DNA
ORGANISM: Homo Sapien
US-10-140-002-363

Query Match 4.0%; Score 42.4; DB 4; Length 1503;
Best Local Similarity 66.3%; Pred. No. 0.41;
Matches 61; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 969 TGACCTGCGCTGGCCCGCCCTACTCTGTGTAATAGATAAGCCCTGCTGTCTGT 1028
DB 1385 TGGCGCTGGCGCTGGGACACCTCTCTGCGCAGAGGCAATAAAGCCGCGGACCT 1444
QY 1029 GGAATAAAGGCTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1060
DB 1445 TGAATAAAGGCTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1476

RESULT 10
US-09-903-603A-220
Sequence 220, Application US/09903603A
Patent No. 6767995
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Deenoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kilavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: GNB.16182C12
CURRENT APPLICATION NUMBER: US/09/903, 603A
CURRENT FILING DATE: 2001-07-11
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143, 048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145, 698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146, 222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944

PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 220
LENGTH: 1503
TYPE: DNA
ORGANISM: Homo sapiens
US-09-903-603A-220

Query Match 4.0%; Score 42.4; DB 4; Length 1503;
Best Local Similarity 66.3%; Pred. No. 0.41;
Matches 61; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 969 TGACCTGCGCTGGCCCGCCCTACTCTGTGTAATAGATAAGCCCTGCTGTCTGT 1028
DB 1385 TGGCGCTGGCGCTGGGACACCTCTCTGCGCAGAGGCAATAAAGCCGCGGACCT 1444
QY 1029 GGAATAAAGGCTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1060
DB 1445 TGAATAAAGGCTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1476

RESULT 11
US-09-311-021-103
Sequence 103, Application US/09311021
Patent No. 6706869
GENERAL INFORMATION:
APPLICANT: Wong, Gordon G.
APPLICANT: Clark, Hilary
APPLICANT: Fechtel, Kim
APPLICANT: Agostino, Michael J.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
FILE REFERENCE: GI 6300-11A
CURRENT APPLICATION NUMBER: US/09/311, 021
CURRENT FILING DATE: 1999-05-13
NUMBER OF SEQ ID NOS: 268
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 103
LENGTH: 977
TYPE: DNA
ORGANISM: Homo sapiens
US-09-311-021-103

Query Match 4.0%; Score 42; DB 4; Length 977;
Best Local Similarity 82.8%; Pred. No. 0.44;
Matches 48; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1003 TAGAATAAAGGCTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1060
DB 919 TGAATAAAGGCTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 976


```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/915,246
; FILING DATE: 19920716
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pochopien, Donald J.
; REGISTRATION NUMBER: 32,167
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312 707-8889
; TELEFAX: 312 707-9155
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1505 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Brassica napus
; STRAIN: Westar
; DEVELOPMENTAL STAGE: Somatic
; TISSUE TYPE: Root
;
US-07-915-246-1

```

Query Match 3.9%; Score 41.8; DB 1; Length 1505;

Best Local Similarity 47.2%; Pred. No. 0.58; Mismatches 142; Indels 0; Gaps 0;

```

QY 552 GAGATCCATTGATGATTAAGTGAACGTTCCGAGAGAGCGGAGAGCCGAGAAAGCGTCTCT 611
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 900 GAGAGAGCTGAGAGACATGATGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 959
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 612 GATGAGAGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 671
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 960 GCGCGTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1019
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 672 TGAGCTGCGCGCTGAGTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 731
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1020 GAGAGTGGCGGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1079
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 732 TTACGAGATCTGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 791
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1080 GAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1139
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 792 AGGCTTGCCTGGAGCGGCTGAGCTGGGCA 820
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1140 GATATGAGAGTGGCGGTGAGTGAAGCA 1168
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 15
US-07-861-458C-37

Sequence 37, Application US/07861458C

Patent No. 6232061

GENERAL INFORMATION:

APPLICANT: Marchionni, Mark Andrew

APPLICANT: Johnson, Carl D.

TITLE OF INVENTION: HOMOLOGY CLONING

NUMBER OF SEQUENCES: 142

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson

STREET: 225 Franklin Street

CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM PS/2 Model 50Z or 55SX

OPERATING SYSTEM: MS-DOS (Version 5.0)

SOFTWARE: WordPerfect (Version 5.1)

```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/861,458C
; FILING DATE: 04/01/92
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1600
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-07-861-458C-37

```

Query Match 3.9%; Score 41.8; DB 3; Length 1600;

Best Local Similarity 71.4%; Pred. No. 0.6; Mismatches 22; Indels 0; Gaps 0;

```

QY 984 CCAGCCCTACTGTGTATATAGATTAAGCGCTGCGTGTCTGTGGAAGAGAGAGAGAG 1043
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1522 CCATGCAAGTGTGTCTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1581
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1044 AAAAAAAAAAAAAAAAAA 1060
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1582 AAAAAAAAAAAAAAAAAA 1598
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: November 22, 2004, 12:08:20
Job time : 209 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 10:12:50 ; Search time 616 Seconds
(without alignments)
9298.443 Million cell updates/sec

Title: US-09-597-920B-1

Perfect score: 1060

Sequence: 1 gacctgcgccttgaggggcc.....aaaaaaaaaaaaaaaaaaaa 1060

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3627888 seqs, 2701811610 residues

Total number of hits satisfying chosen parameters: 7255776

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	949.2	89.5	1224	9	US-09-739-907-13
2	949.2	89.5	1224	11	US-09-938-671-13
3	424	40.0	457	15	US-10-102-524-1371
4	416.8	39.3	528	16	US-10-641-643-565
5	271.6	25.6	319	9	US-09-796-692-5333
6	271.6	25.6	319	14	US-10-040-862-5333
7	271.6	25.6	319	16	US-10-057-475B-5333
8	271.6	25.6	319	16	US-10-154-884B-5333
9	271.6	25.6	319	17	US-10-764-324-5333
C 10	208	19.6	421	9	US-09-796-692-5582
C 11	208	19.6	421	14	US-10-040-862-5582
C 12	208	19.6	421	16	US-10-057-475B-5582

C 13	208	19.6	421	16	US-10-154-884B-5582	Sequence 5582, Ap
C 14	208	19.6	421	17	US-10-764-324-5582	Sequence 5582, Ap
C 15	205.2	19.4	400	16	US-10-242-535A-9083	Sequence 9083, Ap
C 16	205.2	19.4	400	16	US-10-085-783A-9083	Sequence 9083, Ap
C 17	117	11.0	117	9	US-09-796-692-5736	Sequence 5736, Ap
C 18	117	11.0	117	14	US-10-040-862-5736	Sequence 5736, Ap
C 19	117	11.0	117	16	US-10-057-475B-5736	Sequence 5736, Ap
C 20	117	11.0	117	16	US-10-154-884B-5736	Sequence 5736, Ap
C 21	117	11.0	117	17	US-10-764-324-5736	Sequence 5736, Ap
C 22	49.6	4.7	2968	9	US-09-925-301-584	Sequence 584, App
C 23	49.6	4.7	2968	15	US-10-106-698-1500	Sequence 1500, Ap
C 24	48.8	4.6	500	18	US-10-425-115-83126	Sequence 83126, A
C 25	47.4	4.5	1419	17	US-10-437-963-25990	Sequence 25990, A
C 26	47.4	4.5	3196	9	US-09-782-980-50	Sequence 50, Appl
C 27	47.4	4.5	3196	17	US-10-806-018-50	Sequence 50, Appl
C 28	46	4.3	477	10	US-09-918-995-10907	Sequence 10907, A
C 29	45.8	4.3	2022	17	US-10-437-963-21276	Sequence 21276, A
C 30	45.6	4.3	468	10	US-09-918-995-21830	Sequence 21830, A
C 31	45	4.2	496	16	US-10-240-425-154	Sequence 154, App
C 32	45	4.2	777	14	US-10-184-644-348	Sequence 348, App
C 33	45	4.2	777	14	US-10-184-634-348	Sequence 348, App
C 34	44.8	4.2	173	9	US-09-834-975-589	Sequence 589, App
C 35	44.8	4.2	189	9	US-09-834-975-390	Sequence 390, App
C 36	44.8	4.2	848	9	US-09-823-245A-186	Sequence 186, App
C 37	44.6	4.2	424	9	US-09-867-701-1317	Sequence 1317, Ap
C 38	44.6	4.2	472	10	US-09-918-995-16588	Sequence 16588, A
C 39	44.4	4.2	910	14	US-10-198-846-6483	Sequence 6483, A
C 40	44.2	4.2	178	17	US-10-437-963-23303	Sequence 23303, A
C 41	44.2	4.2	512	18	US-10-425-115-131750	Sequence 131750, A
C 42	44.2	4.2	1338	14	US-10-097-065-19	Sequence 19, Appl
C 43	44.2	4.2	1338	15	US-10-372-876-19	Sequence 19, Appl
C 44	44.2	4.2	1537	10	US-09-820-788-1	Sequence 1, Appl
C 45	44.2	4.2	1677	16	US-10-264-237-401	Sequence 401, App

ALIGNMENTS

RESULT 1
US-09-739-907-13
; Sequence 13, Application US/09739907
; Patent No. US20010012869A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2022P1
; CURRENT APPLICATION NUMBER: US/09/739,907
; CURRENT FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 09/348,457
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: 60/070,567
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,692
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1224
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1205)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-739-907-13
Query Match 89.5%; Score 949.2; DB 9; Length 1224;
Best Local Similarity 98.9%; Pred. No. 1.5e-261;
Matches 1039; Conservative 0; Mismatches 4; Indels 8; Gaps 8;

OY	1	GACTGTGCCCTTGAAGGGGCTTAAGGGGGTCAAGACACCTGTCTCCAGACTCCCTCGAAGT	60
Db	181	GACTGTGCCCTTGAAGGGGCTTAAGGGGGTCAAGACACCTGTCTCCAGACTCCCTCGAAGT	240
OY	61	GAGGAGGCGCATCTGTGTCCCTGTGCGTGTGGGGCTCTGTGCTGTGCCATCTTGCGCATG	120
Db	241	GAGGAGGCGCATCTGTGTCCCTGTGCGTGTGGGGCTCTGTGCTGTGCCATCTTGCGCATG	300
OY	121	TTGATGTGGACATGTGTGTGACATGCGCACAGATGCGCAGGGCTCTTACAGACAGACATCTCA	180
Db	301	TTGATGTGGACATGTGTGTGACATGCGCACAGATGCGCAGGGCTCTTACAGACAGACATCTCA	360
OY	181	GATAGTTTGTATTC -AAGGGGATCCAGTTCAAAAGGCTCAACAAGTGTGCCCTGTGACC	239
Db	361	GATAGTTTGTATTCAAAGGGGATCCAGTTCAAAAGGCTCAACAAGTGTGCCCTGTGACC	420
OY	240	ACCTGTCCATCCCACTGTATCACTCTTACCACCCCTGTAGCCAGACAGACTGTCTCCCAT	299
Db	421	ACCTGTCCATCCCACTGTATCACTCTTACCACCCCTGTAGCCAGACAGACTGTCTCCCAT	480
OY	300	CCCAAGATCCCCGCGAGCCCTTTGGGGGGCTCCCAACGCGCATTTTCCGGCGGATTC	359
Db	481	CCCAAGATCCCCGCGAGCCCTTTGGGGGGCTCCCAACGCGCATTTTCCGGCGGATTC	540
OY	360	TGATGTGCCAAACAGTGTGTGGAGCTACAGAAACAGGAAACAGCTGTGAGATGACGA	419
Db	541	TGATGTGCCAAACAGTGTGTGGAGCTACAGAAACAGGAAACAGCTGTGAGATGACGA	599
OY	420	TGAGGATAGGAGACGATACATCAACCCAGGCTACCTGTGTGTCTCTGTACAGACACCC	479
Db	600	TGA -GATAGGACGATACATCAACCCAGGCTACCTGTGTGTCTCTGTACAGACACCC	658
OY	480	GCGCATAGACACTGTGCCCCATCACTCTGTACACACCCCTGTGCATCCGAGACAG	539
Db	659	GCGCATAGACACTGTGCCCCATCACTCTGTACACACCCCTGTGCATCCGAGACAG	717
OY	540	TGCTTTCTCCATGATGTCCATTGATGATTAAGTAAAGTTCCGAGACCGGGGAGAGGCG	599
Db	718	TGCTTTCTCCATGATGTCCATTGATGATTAAGTAAAGTTCCGAGACCGGGGAGAGGCG	777
OY	600	AGAAAGCTCTCTGATGTGCGACGCGGAGATATGTGATGTGCCAGAACTGCATCTCTG	659
Db	778	AGAAAGCTCTCTGATGTGCGACGCGGAGATATGTGATGTGCCAGAACTGCATCTCTG	837
OY	660	AGCGGCTAAGACTGAGCTGCCGCCCTTAAGTTCCAGAGGACAGAGAAAGTGAAGAA	719
Db	838	AGCGGCTAAGACTGAGCTGCCGCCCTTAAGTTCCAGAGGACAGAGAAAGTGAAGAA	897
OY	720	GGGGGCTCCAGATTACGAGAACTGTGACAGAGCTGAACAGGGGCTGTGAGAGCGGAACTC	779
Db	898	GGGGGCTCCAGATTACGAGAACTGTGACAGAGCTGAACAGGGGCTGTGAGAGCGGAACTC	957
OY	780	TGTCTGTGAACACGAGCTTGTGCTGTGAGACGAGCTGAGCTGTGAGAGTGTGCTGTGG	839
Db	958	TGTCTGTGAACACGAGCTTGTGCTGTGAGACGAGCTGAGCTGTGAGAGTGTGCTGTGG	1017
OY	840	TCCTACACATGGGCTGTGCTGTGCTTCCAGCTGTACAAACAGCTGTGAAGAAATCCCCGTA	899
Db	1018	TCCTACACATGGGCTGTGCTGTGCTTCCAGCTGTACAAACAGCTGTGAAGAAATCCCCGTA	1077
OY	900	ACTTATTTACATTTTGGGGGTTGGGCTGTGTCCCGGAAAGCTGTGACCTTGTGACGCA	959
Db	1078	ACTTATTTACATTTTGGGGGTTGGGCTGTGT -CCCCGAAAGCTGTG -ACCTTCTGACGCA	1135
OY	960	GCTTGAGATGACCTGTGCCCTGTGCGCCAGCCCTTACTGTGTATATGAATTAAGGCTTGC	1019
Db	1136	GCTTGAGATGACCTGTGCGCCAGCCCTTACTGTG -GTATATCAATTAAGGCTTGC	1193
OY	1020	TGTGTCTGTGAAAAAAAAAAAAAAAAAAAA 1050	
Db	1194	TGTGTCTGTGTAAAAAAAAAAAAAAAAAAAA 1224	

```

US-09-938-671-13
RESULT 2
US-09-938-671-13
Sequence 13, Application US/09938671
Publication NO. US20040002066A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 36 Human Secreted Proteins
FILE REFERENCE: P2022P1
CURRENT APPLICATION NUMBER: US/09/938, 671
CURRENT FILING DATE: 2001-08-27
PRIOR APPLICATION NUMBER: 09/348,457
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: 60/070,567
PRIOR FILING DATE: 1998-01-07
PRIOR APPLICATION NUMBER: 60/070,692
PRIOR FILING DATE: 1998-01-07
PRIOR APPLICATION NUMBER: 60/070,704
PRIOR FILING DATE: 1998-01-07
PRIOR APPLICATION NUMBER: 60/070,658
PRIOR FILING DATE: 1998-01-07
NUMBER OF SEQ ID NOS: 196
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 13
LENGTH: 1224
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (1205)
OTHER INFORMATION: n equals a,t,g, or c
US-09-938-671-13
```

Query Match	Score	DB	Length
89.5%	949.2	DB 11	1224

Best Local Similarity 98.9%; Pred. No. 1.5e-261;
Matches 1039; Conservative 0; Mismatches 4; Indels 8; Gaps 8;

OY	1	GACTGTGGCCCTTGAAGGGGCTTAAGGGGTTCACACAGCCTGCTCCGAGCTCCCTCGAGATG	60
Db	181	GACTGTGCCCTTGAAGGGGCTTAAGGGGTTCACACAGCCTGCTCCGAGCTCCCTCGAGATG	240
OY	61	GAGAGGCGCATCTGTGTTCCCTCGCGTGTGGGGCTCTGTGCTTACCATCTTGCCATG	120
Db	241	GAGAGGCGCATCTGTGTTCCCTCGCGTGTGGGGCTCTGTGCTTACCATCTTGCCATG	300
OY	121	TTGATGAGCACTGTGTGTGCTACTGCCACAGACTGCCAGGCTCTCTACGACAGACATCTTCA	180
Db	301	TTGATGAGCACTGTGTGTGCTACTGCCACAGACTGCCAGGCTCTCTACGACAGACATCTTCA	360
OY	181	GATGATTGTGTATCCAAAGGGGATTCAGATTCAAAGGCTCTACACAGTTGGCCCCCTGGCC	239
Db	361	GATGATTGTGTATCCAAAGGGGATTCAGATTCAAAGGCTCTACACAGTTGGCCCCCTGGCC	420
OY	240	AACGTGCCATCCACCTGTCACTCTCACTCCACACCCCTGAGCGACAGACCTGTCCCAT	299
Db	421	AACGTGCCATCCACCTGTCACTCTCACTCCACACCCCTGAGCGACAGACCTGTCCCAT	480
OY	300	CCCAAGATTCGCCGAGCCCTTGGGGGCTCCACACCGACGCAATCTTCCGGCGGGATTTC	359
Db	481	CCCAAGATTCGCCGAGCCCTTGGGGGCTCCACACCGACGCAATCTTCCGGCGGGATTTC	540
OY	360	TGATGTGCCAACAGTGTGGCGAGCTACGAGAACGAGGAACCAAGCTGTAGAGATGCAGA	419
Db	541	TGATGTGCCAACAGTGTGGCGAGCTACGAGAACGAGGAACCAAGCTGTAGAGATGCAGA	599
OY	420	TGAGGATGAGGACGACTATCAACAACCAAGGCTACTGTGTGTGCTTCTTGACAGTACCCC	479
Db	600	TGA-GATAGGAGCGACTATCAACAACCAAGGCTACTGTGTGTGCTTCTTGACAGTACCCC	658
OY	480	GGCCACTTGAAGCACTGTGCCCCATGACGCTCGACACTGACACCCCTGGCATCTCGAGACAG	539
Db	659	GGCCACTTGAAGCACTGTGCCCCATGACGCTCGACACTGACACCCCTGGCATCTCGAGACAG	717

Qy 540 TGCCTTCATGAGTCATGATGATTAACGTGACGCTCCGAGAGCGGAGAGCGC 599
Db 718 TGCCTTCATGAGTCATGATGATTAACGTGACGCTCCGAGAGCGGAGAGCGC 777
Qy 600 AGAAGCGCTTCGATGAGCAGCGGAGATGATGATGTCCTCCAGAACTGCATCTCG 659
Db 778 AGAAGCGCTTCGATGAGCAGCGGAGATGATGATGTCCTCCAGAACTGCATCTCG 837
Qy 660 AGCGGCTAAGACTGAGCGCTGCGGCTGATGTCCTCCAGAGAGGAGAGGAGAG 719
Db 838 AGCGGCTAAGACTGAGCGCTGCGGCTGATGTCCTCCAGAGAGGAGAGGAGAG 897
Qy 720 GGGGCTCCAGATTAAGAGAACTGACAGAGCTGAATGAGGCTGTGAGAGCGGATC 779
Db 898 GGGGCTCCAGATTAAGAGAACTGACAGAGCTGAATGAGGCTGTGAGAGCGGATC 957
Qy 780 TGTCTGGAACGAGGCTTGCCTGAGAGCGCTGAGAGCTGAGAGGAGGCTGTGAG 839
Db 958 TGTCTGGAACGAGGCTTGCCTGAGAGCGCTGAGAGCTGAGAGGAGGCTGTGAG 1017
Qy 840 TCCCTCAATGAGCGCTGCGGCTGCTGAGCTGACCAAGCGCTGAGAAATCCCGGCTA 899
Db 1018 TCCCTCAATGAGCGCTGCGGCTGCTGAGCTGACCAAGCGCTGAGAAATCCCGGCTA 1077
Qy 900 ACTTATTATCACTTTGGGGGTTGCGGCTGTGTCCTCCGAAAGCTGTGACCTTGTGACCA 959
Db 1078 ACTTATTATCACTTTGGGGGTTGCGGCTGTGTCCTCCGAAAGCTGTGACCTTGTGACCA 1135
Qy 960 GCTGAGAGATGACCTGCGGCTGCGGCTGCTGATGATGATGATGATGATGATGATG 1019
Db 1136 GCTGAGAGATGACCTGCGGCTGCGGCTGCTGATGATGATGATGATGATGATGATG 1193
Qy 1020 TGTGCTGTGAGAAAAAATGAAAAAATGAAAAA 1050
Db 1194 TGTGCTGTGAGAAAAAATGAAAAAATGAAAAA 1224

RESULT 3

US-10-102-524-1371
; Sequence 1371, Application US/10102524
; Publication No. US20030109434A1
; GENERAL INFORMATION:
; APPLICANT: Mannion, Paul A.
; APPLICANT: Algate, Paul A.
; APPLICANT: Gaiger, Alexander
; APPLICANT: Gordon, Brian
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; FILE REFERENCE: 210121.572
; CURRENT APPLICATION NUMBER: US/10/102.524
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 1863
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1371
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 449_
; OTHER INFORMATION: n = A,T,C or G
US-10-102-524-1371

Query Match 40.0%; Score 424; DB 15; Length 457;
Best Local Similarity 99.8%; Pred. No. 3.6e-11;
Matches 424; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 397 GAACCAAGCTTGTGAGATGAGATGAGATGAGACGCTATCAACACCGGCTACTG 456
Db 33 GAACCAAGCTTGTGAGATGAGATGAGATGAGACGCTATCAACACCGGCTACTG 92
Qy 457 GTGTGTCTTGTGACAGACCCGGGCACTAGCACTGTGCCCCATCAAGTCTGCACTC 516

Db 93 GTGTGTCTTGTGACAGACCCGGGCACTAGCACTGTGCCCCATCAAGTCTGCACTC 152
Qy 517 AGCAACCCCTGGCATCCAGAGACAGTGCCTTCTCCATGAGATGCATGATGATTAAGTGAAC 576
Db 153 AGCAACCCCTGGCATCCAGAGACAGTGCCTTCTCCATGAGATGCATGATGATTAAGTGAAC 212
Qy 577 GTTCCGAGAGCGGGGAGAGGCGCAGAAAGCTGTGATGAGAGCGGAGATGATGAAT 636
Db 213 GTTCCGAGAGCGGGGAGAGGCGCAGAAAGCTGTGATGAGAGCGGAGATGATGAAT 272
Qy 637 GTGTCCAGAACTGCATCTCTGAGCGGCTTAAGACTGAGCTGCGGCTGAGTTCCAG 696
Db 273 GTGTCCAGAACTGCATCTCTGAGCGGCTTAAGACTGAGCTGCGGCTGAGTTCCAG 332
Qy 697 GAGGCAAGAGAGTGTGAGAGAGAGGCGGCTCCAGATTTCAGAGATCTCAGAGAGCTGAAC 756
Db 333 GAGGCAAGAGAGTGTGAGAGAGAGGCGGCTCCAGATTTCAGAGATCTCAGAGAGCTGAAC 392
Qy 757 TGAGGCGCTGTGAGAGCGGAGTGTCTGTGAAACAGGCTTGCGGAGCGGCTGAGCTG 816
Db 393 TGAGGCGCTGTGAGAGCGGAGTGTCTGTGAAACAGGCTTGCGGAGCGGCTGAGCTG 452
Qy 817 GGCAG 821
Db 453 GGCAG 457

RESULT 4

US-10-641-643-565
; Sequence 565, Application US/10641643
; Publication No. US20040077003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
; GENE EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/641.643
; FILING DATE: 14-Aug-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 565:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 528 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: HIPNOTO1

```

; CLONE: 240885
; SEQUENCE DESCRIPTION: SEQ ID NO: 565
US-10-641-643-565

```

Query Match	33.3%	Score	416.8	DB16	Length	528			
Best Local Similarity	99.3%	Pred	No. 4.4e-109						
Matches	429	Conservative	0	Mismatches	2	Indels	1	Gaps	1

QY	599	CAGAAAGCCTCTCTGGATGGCAGCCGGAGATATGTAAATGTGTCCAGAGAACTGCATCCTGG	658
Db	1	CAGAAAGCCTCTCTGGATGGCAGCCGGAGATATGTAAATGTGTCCAGAGAACTGCATCCTGG	60
QY	659	GAGCGGCTTAAGACTGAGCCTCGCCGCTTGAGTTCCACAGAGCAGAGAAATGAGAGAG	718
Db	61	GAGCGGCTTAAGACTGAGCCTCGCCGCTTGAGTTCCACAGAGCAGAGAAATGAGAGAG	120
QY	719	AGGGGGCTCCAGATTACAGAAATCTGACAGAGCTGAACCTGAGGGGCTGTGAGAGCCGAGT	778
Db	121	AGGGGGCTCCAGATTACAGAAATCTGACAGAGCTGAACCTGAGGGGCTGTGAGAGCCGAGT	180
QY	779	CTGTCTTGGAACACAGGCTTTGCCTGGG-AAGGCTGAGCTTGAGCAGCTGGAAGTGAGCTTGG	837
Db	181	CTGTCTTGGAACACAGGCTTTGCCTGGGAAGGCTGAGCTTGAGCAGCTGGAAGTGAGCTTGG	240
QY	838	GGTCTCTACATAGGGGCTCTGCGCTTGTCTCAAGCTGACACAAACAGCTTGAGAAATCCCCCGG	897
Db	241	GGTCTCTACATAGGGGCTCTGCGCTTGTCTCAAGCTGACACAAACAGCTTGAGAAATCCCCCGG	300
QY	898	TAACTATTATCACTTTGAGGGTTGGGCTGTGTGCCCGGAACGCTGTGACACTTCTGAGC	957
Db	301	TAACTATTATCACTTTGAGGGTTGGGCTGTGTGCCCGGAACGCTGTGACACTTCTGAGC	360
QY	958	CAGCCTGGAATGACCTGSCCTTGAGGCCCAAGCCCTAATCTGTGTAAATGAATTAAGGCGTG	1017
Db	361	CAGCCTGGAATGACCTGSCCTTGAGGCCCAAGCCCTAATCTGTGTAAATGAATTAAGGCGTG	420
QY	1018	CGTGTGTCTGTG 1029	
Db	421	CGTGTGTCTATG 432	

```

RESULT 5
US-09-796-692-5333
: Sequence 5333, Application US/097966692
: Publication No. US20020198362A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
FILE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
FILE REFERENCE: 2077_001200
CURRENT APPLICATION NUMBER: US/09/796,692
CURRENT FILING DATE: 2001-03-01
PRIORITY APPLICATION NUMBER: 60/186,126
PRIORITY FILING DATE: 2000-03-01
PRIORITY APPLICATION NUMBER: 60/190,479
PRIORITY FILING DATE: 2000-03-17
PRIORITY APPLICATION NUMBER: 60/200,545
PRIORITY FILING DATE: 2000-04-27
PRIORITY APPLICATION NUMBER: 60/200,303
PRIORITY FILING DATE: 2000-04-28
PRIORITY APPLICATION NUMBER: 60/200,779
PRIORITY FILING DATE: 2000-04-28
PRIORITY APPLICATION NUMBER: 60/200,999
PRIORITY FILING DATE: 2000-05-01
PRIORITY APPLICATION NUMBER: 60/202,084
PRIORITY FILING DATE: 2000-05-04
PRIORITY APPLICATION NUMBER: 60/206,201
PRIORITY FILING DATE: 2000-05-22
PRIORITY APPLICATION NUMBER: 60/218,950
PRIORITY FILING DATE: 2000-07-14
PRIORITY APPLICATION NUMBER: 60/222,903

```

```

; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5333
; LENGTH: 319
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-5333

```

Query Match	25.6%	Score 271.6	DB 9	Length 319
Best Local Similarity	98.6%	Pred. No. 1.6e-67		
Matches 274	Conservative 0	Mismatches 4	Indels 0	Gaps 0

Accession	Sequence	Position
QY	TTCTCCATGAGATGCATATGATGATTACGTAAACCTCCGAGAGACGGGGAGAGCGCAGAA	603
Db	TCCACAGTGGAGTCCATTGATGATTACGTAAACCTCCGAGAGACGGGGAGAGCGCAGAA	101
QY	GGCTCTCTGGATGCGACGCCGGAGATATGTGATGTCTCCAGGAACTGCATCTTGGACCG	663
Db	GGCTCTCTGGATGCGACGCCGGAGATATGTGATGTCTCCAGGAACTGCATCTTGGACCG	161
QY	GCTAAGACTGAGCGCTGCGCCCTCTAGTGTCCCAAGAGGACAGAGAAAGTGGAGGAAGAGGG	723
Db	GCTAAGACTGAGCGCTGCGCCCTCTAGTGTCCCAAGAGGACAGAGAAAGTGGAGGAAGAGGG	221
QY	GCTCCAGATTACAGAAATCTGACAGAGCTGAACTGAGGGCGTGTGAGAGCCGAGTCTGTCT	783
Db	GCTCCAGATTACAGAAATCTGACAGAGCTGAACTGAGGGCGTGTGAGAGCCGAGTCTGTCT	281
QY	CTGGACACAGGCTTGCTGGGACGGCTGAGAGCTGGGGCAG	821
Db	CTGGACACAGGCTTGCTGGGACGGCTGAGAGCTGGGGCAG	319

RESULT 6
US-10-040-862-5333
Sequence 5333, Application US/10040862
Publication No. US20030078396A1
GENERAL INFORMATION:
APPLICANT: Galger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Retter, Marc
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
of
TITLE OF INVENTION: Hematological Malignancies
FILE REFERENCE: 014058-013520US
CURRENT APPLICATION NUMBER: US/10/040,862
CURRENT FILING DATE: 2001-11-06
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/130,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03

;; PRIOR APPLICATION NUMBER: US 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: US 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; PRIOR APPLICATION NUMBER: US 09/796,692
;; PRIOR FILING DATE: 2001-03-01
;; NUMBER OF SEQ ID NOS: 10467
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO: 5333
;; LENGTH: 319
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-040-862-5333

Query Match 25.6%; Score 271.6; DB 14; Length 319;
Best Local Similarity 98.6%; Pred. No. 1.6e-67;
Matches 274; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 544 TTCTCATGAGTCCATTGATGATTACGTAACGTTCCGAGAGCGGGAGAGCCAGAA 603
Db 42 TCCACAGTGGAGTCCATTGATGATTACGTAACGTTCCGAGAGCGGGAGAGCCAGAA 101
Qy 604 GCGTCTGATGAGCGAGCGGGAGTATGTGTAATGTGTCCAGAACTGCATCTGAGCG 663
Db 102 GCGTCTGATGAGCGAGCGGGAGTATGTGTAATGTGTCCAGAACTGCATCTGAGCG 161
Qy 664 GCTAAGACTGAGCGTCCGCGCCCTGAGTCCAGAGGCGAGAACTGAGAGAAAGGGG 723
Db 162 GCTAAGACTGAGCGTCCGCGCCCTGAGTCCAGAGGCGAGAACTGAGAGAAAGGGG 221
Qy 724 GCTCAGATTGAGAACTGAGAGGCTGAACTGAGAGGCGCTGTGAGAGCGGAGTGTGC 783
Db 222 GCTCAGATTGAGAACTGAGAGGCTGAACTGAGAGGCGCTGTGAGAGCGGAGTGTGC 281
Qy 784 CTGGAACCAAGCGCTTGCCTGGAGCGGCTGAGCTGGGCGAG 821
Db 282 CTGGAACCAAGCGCTTGCCTGGAGCGGCTGAGCTGGGCGAG 319

RESULT 7
US-10-057-475B-5333
;; Sequence 5333, Application US/10057475B
;; Publication No. US20040002068A1
;; GENERAL INFORMATION:
;; APPLICANT: Gaiger, Alexander
;; APPLICANT: Algate, Paul A.
;; APPLICANT: Mannion, Jane
;; APPLICANT: Clapper, Jonathan David
;; APPLICANT: Wang, Aijun
;; APPLICANT: Ordenez, Nadia
;; APPLICANT: Carter, Lauren
;; APPLICANT: McNeill, Patricia Dianne
;; TITLE OF INVENTION: Corixa Corporation
;; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
;; FILE REFERENCE: 014058-014402US
;; CURRENT APPLICATION NUMBER: US/10/057,475B
;; PRIOR FILING DATE: 2002-01-22
;; PRIOR APPLICATION NUMBER: US 60/186,126
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: US 60/190,479
;; PRIOR FILING DATE: 2000-03-17
;; PRIOR APPLICATION NUMBER: US 60/200,545
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: US 60/200,303
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: US 60/200,779
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: US 60/200,999
;; PRIOR FILING DATE: 2000-05-01
;; PRIOR APPLICATION NUMBER: US 60/202,084
;; PRIOR FILING DATE: 2000-05-04
;; PRIOR APPLICATION NUMBER: US 60/206,201

;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: US 60/218,950
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: US 60/222,903
;; PRIOR FILING DATE: 2000-08-03
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 10979
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO: 5333
;; LENGTH: 319
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-057-475B-5333

Query Match 25.6%; Score 271.6; DB 16; Length 319;
Best Local Similarity 98.6%; Pred. No. 1.6e-67;
Matches 274; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 544 TTCTCATGAGTCCATTGATGATTACGTAACGTTCCGAGAGCGGGAGAGCCAGAA 603
Db 42 TCCACAGTGGAGTCCATTGATGATTACGTAACGTTCCGAGAGCGGGAGAGCCAGAA 101
Qy 604 GCGTCTGATGAGCGAGCGGGAGTATGTGTAATGTGTCCAGAACTGCATCTGAGCG 663
Db 102 GCGTCTGATGAGCGAGCGGGAGTATGTGTAATGTGTCCAGAACTGCATCTGAGCG 161
Qy 664 GCTAAGACTGAGCGTCCGCGCCCTGAGTCCAGAGGCGAGAACTGAGAGAAAGGGG 723
Db 162 GCTAAGACTGAGCGTCCGCGCCCTGAGTCCAGAGGCGAGAACTGAGAGAAAGGGG 221
Qy 724 GCTCAGATTGAGAACTGAGAGGCTGAACTGAGAGGCGCTGTGAGAGCGGAGTGTGC 783
Db 222 GCTCAGATTGAGAACTGAGAGGCTGAACTGAGAGGCGCTGTGAGAGCGGAGTGTGC 281
Qy 784 CTGGAACCAAGCGCTTGCCTGGAGCGGCTGAGCTGGGCGAG 821
Db 282 CTGGAACCAAGCGCTTGCCTGGAGCGGCTGAGCTGGGCGAG 319

RESULT 8
US-10-154-884B-5333
;; Sequence 5333, Application US/10154884B
;; Publication No. US20040005561A1
;; GENERAL INFORMATION:
;; APPLICANT: Gaiger, Alexander
;; APPLICANT: Algate, Paul A.
;; APPLICANT: Mannion, Jane
;; APPLICANT: Retter, Marc W.
;; APPLICANT: Corixa Corporation
;; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
;; FILE REFERENCE: 014058-013521US
;; CURRENT APPLICATION NUMBER: US/10/154,884B
;; PRIOR FILING DATE: 2002-05-23
;; PRIOR APPLICATION NUMBER: US 60/186,126
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: US 60/190,479
;; PRIOR FILING DATE: 2000-03-17
;; PRIOR APPLICATION NUMBER: US 60/200,545
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: US 60/200,303
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: US 60/200,779
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: US 60/200,999
;; PRIOR FILING DATE: 2000-05-01
;; PRIOR APPLICATION NUMBER: US 60/202,084
;; PRIOR FILING DATE: 2000-05-04
;; PRIOR APPLICATION NUMBER: US 60/206,201
;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: US 60/218,950
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: US 60/222,903

```

; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5333
; LENGTH: 319
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-984B-5333

Query Match      25.6%; Score 271.6; DB 16; Length 319;
Best Local Similarity 98.6%; Pred. No. 1.6e-67;
Matches 274; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 544 TTCTCCATGGAGTCATTGATGATTACGTGAACGTTCCGAGAGCGGGAGCGGAGAA 603
Db 42 TCACAGTGGAGTCATTGATGATTACGTGAACGTTCCGAGAGCGGGAGCGGAGAA 101
QY 604 GCGTCTGTGATGGACCGCGGAGTATGTGATGTCTCCAGAACTGCATCTTGAGCG 663
Db 102 GCGTCTGTGATGGACCGCGGAGTATGTGATGTCTCCAGAACTGCATCTTGAGCG 161
QY 664 GCTTAAGCTGAGCCTGCGCCCTGAGTTCCTCCAGAGGCAAGAGAGAGAGGG 723
Db 162 GCTTAAGCTGAGCCTGCGCCCTGAGTTCCTCCAGAGGCAAGAGAGAGAGGG 221
QY 724 GCTTCAATTACAGAACTGTGACAGAGCTGAAGAGGCGCTGTGAGAGCCGAGTCTGTC 783
Db 222 GCTTCAATTACAGAACTGTGACAGAGCTGAAGAGGCGCTGTGAGAGCCGAGTCTGTC 281
QY 784 CTGGAACCAAGGCTTGCTGGGAGCGCTGAGCTGGGCGAG 821
Db 282 CTGGAACCAAGGCTTGCTGGGAGCGCTGAGCTGGGCGAG 319

RESULT 9
US-10-764-324-5333
; Sequence 5333, Application US/10764324
; Publication No. US2004017539A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/764,324
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
```

```

; SEQ ID NO 5333
; LENGTH: 319
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-764-324-5333

Query Match      25.6%; Score 271.6; DB 17; Length 319;
Best Local Similarity 98.6%; Pred. No. 1.6e-67;
Matches 274; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 544 TTCTCCATGGAGTCATTGATGATTACGTGAACGTTCCGAGAGCGGGAGCGGAGAA 603
Db 42 TCACAGTGGAGTCATTGATGATTACGTGAACGTTCCGAGAGCGGGAGCGGAGAA 101
QY 604 GCGTCTGTGATGGACCGCGGAGTATGTGATGTCTCCAGAACTGCATCTTGAGCG 663
Db 102 GCGTCTGTGATGGACCGCGGAGTATGTGATGTCTCCAGAACTGCATCTTGAGCG 161
QY 664 GCTTAAGCTGAGCCTGCGCCCTGAGTTCCTCCAGAGGCAAGAGAGAGAGGG 723
Db 162 GCTTAAGCTGAGCCTGCGCCCTGAGTTCCTCCAGAGGCAAGAGAGAGAGGG 221
QY 724 GCTTCAATTACAGAACTGTGACAGAGCTGAAGAGGCGCTGTGAGAGCCGAGTCTGTC 783
Db 222 GCTTCAATTACAGAACTGTGACAGAGCTGAAGAGGCGCTGTGAGAGCCGAGTCTGTC 281
QY 784 CTGGAACCAAGGCTTGCTGGGAGCGCTGAGCTGGGCGAG 821
Db 282 CTGGAACCAAGGCTTGCTGGGAGCGCTGAGCTGGGCGAG 319

RESULT 10
US-09-796-692-5582/c
; Sequence 5582, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077,001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5582
; LENGTH: 421
; TYPE: DNA
```

ORGANISM: Homo sapiens
US-09-796-692-5582

Query Match 19.6%; Score 208; DB 9; Length 421;

Best Local Similarity 100.0%; Pred. No. 2.8e-49; Mismatches 0; Indels 0; Gaps 0;

Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 822 CTGGAAGTGGCTCTGGGGTCTCAGATGGCGCTGCGCTTCTCAGCTGACAACAGC 881
DB 421 CTGGAAGTGGCTCTGGGGTCTCAGATGGCGCTGCGCTTCTCAGCTGACAACAGC 362
QY 882 CTGGAAGTGGCTCTGGGGTCTCAGATGGCGCTGCGCTTCTCAGCTGACAACAGC 941
DB 361 CTGGAAGTGGCTCTGGGGTCTCAGATGGCGCTGCGCTTCTCAGCTGACAACAGC 302
QY 942 TCTGACCTTCTGACGACGCTGAGAAATGACCTGCGCTGCGCCAGCCCTACTCTGTGTA 1001
DB 301 TCTGACCTTCTGACGACGCTGAGAAATGACCTGCGCTGCGCCAGCCCTACTCTGTGTA 242
QY 1002 ATGAAATAAGGCTGCGTGTCTGTG 1029
DB 241 ATGAAATAAGGCTGCGTGTCTGTG 214

RESULT 11

US-10-040-862-5582/c

Sequence 5582, Application US/10040862

Publication No. US20030078396A1

GENERAL INFORMATION:

APPLICANT: Gaiger, Alexander

APPLICANT: Algate, Paul A.

APPLICANT: Mannion, Jane

APPLICANT: Retter, Marc

APPLICANT: Corixa Corporation

TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy

FILE OF INVENTION: Hematological Malignancies

FILE REFERENCE: 014058-013520US

CURRENT APPLICATION NUMBER: US/10/040,862

PRIOR FILING DATE: 2001-11-06

PRIOR APPLICATION NUMBER: US 60/186,126

PRIOR FILING DATE: 2000-03-01

PRIOR APPLICATION NUMBER: US 60/190,479

PRIOR FILING DATE: 2000-03-17

PRIOR APPLICATION NUMBER: US 60/200,545

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: US 60/200,303

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: US 60/200,779

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: US 60/200,999

PRIOR FILING DATE: 2000-05-01

PRIOR APPLICATION NUMBER: US 60/202,084

PRIOR FILING DATE: 2000-05-04

PRIOR APPLICATION NUMBER: US 60/206,201

PRIOR FILING DATE: 2000-05-22

PRIOR APPLICATION NUMBER: US 60/218,950

PRIOR FILING DATE: 2000-07-14

PRIOR APPLICATION NUMBER: US 60/222,903

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: US 60/223,416

PRIOR FILING DATE: 2000-08-04

PRIOR APPLICATION NUMBER: US 60/223,378

PRIOR FILING DATE: 2000-08-07

PRIOR APPLICATION NUMBER: US 09/796,692

PRIOR FILING DATE: 2001-03-01

NUMBER OF SEQ ID NOS: 10467

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 5582

LENGTH: 421

TYPE: DNA

ORGANISM: Homo sapiens

US-10-040-862-5582

Query Match 19.6%; Score 208; DB 14; Length 421;

Best Local Similarity 100.0%; Pred. No. 2.8e-49; Mismatches 0; Indels 0; Gaps 0;

Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 822 CTGGAAGTGGCTCTGGGGTCTCAGATGGCGCTGCGCTTCTCAGCTGACAACAGC 881
DB 421 CTGGAAGTGGCTCTGGGGTCTCAGATGGCGCTGCGCTTCTCAGCTGACAACAGC 362
QY 882 CTGGAAGTGGCTCTGGGGTCTCAGATGGCGCTGCGCTTCTCAGCTGACAACAGC 941
DB 361 CTGGAAGTGGCTCTGGGGTCTCAGATGGCGCTGCGCTTCTCAGCTGACAACAGC 302
QY 942 TCTGACCTTCTGACGACGCTGAGAAATGACCTGCGCTGCGCCAGCCCTACTCTGTGTA 1001
DB 301 TCTGACCTTCTGACGACGCTGAGAAATGACCTGCGCTGCGCCAGCCCTACTCTGTGTA 242
QY 1002 ATGAAATAAGGCTGCGTGTCTGTG 1029
DB 241 ATGAAATAAGGCTGCGTGTCTGTG 214

RESULT 12

US-10-057-475B-5582/c

Sequence 5582, Application US/10057475B

Publication No. US2004002068A1

GENERAL INFORMATION:

APPLICANT: Gaiger, Alexander

APPLICANT: Algate, Paul A.

APPLICANT: Mannion, Jane

APPLICANT: Clapper, Jonathan David

APPLICANT: Wang, Aijun

APPLICANT: Ordenez, Nadia

APPLICANT: Carter, Lauren

APPLICANT: McNeill, Patricia Dianne

APPLICANT: Corixa Corporation

TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy

FILE OF INVENTION: Hematological Malignancies

FILE REFERENCE: 014058-014402US

CURRENT APPLICATION NUMBER: US/10/057,475B

PRIOR FILING DATE: 2002-01-22

PRIOR APPLICATION NUMBER: US 60/186,126

PRIOR FILING DATE: 2000-03-01

PRIOR APPLICATION NUMBER: US 60/190,479

PRIOR FILING DATE: 2000-03-17

PRIOR APPLICATION NUMBER: US 60/200,545

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: US 60/200,303

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: US 60/200,779

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: US 60/200,999

PRIOR FILING DATE: 2000-05-01

PRIOR APPLICATION NUMBER: US 60/202,084

PRIOR FILING DATE: 2000-05-04

PRIOR APPLICATION NUMBER: US 60/206,201

PRIOR FILING DATE: 2000-05-22

PRIOR APPLICATION NUMBER: US 60/218,950

PRIOR FILING DATE: 2000-07-14

PRIOR APPLICATION NUMBER: US 60/222,903

PRIOR FILING DATE: 2000-08-03

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 10979

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 5582

LENGTH: 421

TYPE: DNA

ORGANISM: Homo sapiens

US-10-057-475B-5582

Query Match 19.6%; Score 208; DB 16; Length 421;

Best Local Similarity 100.0%; Pred. No. 2.8e-49; Mismatches 0; Indels 0; Gaps 0;

Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 822 CTGGAAGTGGCTCTGGGGTCTCTACATGCGCTCTGCTCCCTTGTCTCCAGCTGACACAGC 881
DB 421 CTGGAAGTGGCTCTGGGGTCTCTACATGCGCTCTGCTCCCTTGTCTCCAGCTGACACAGC 362
QY 882 CTGGAAGTGGCTCTGGGGTCTCTACATGCGCTCTGCTCCCTTGTCTCCAGCTGACACAGC 941
DB 361 CTGGAAGTGGCTCTGGGGTCTCTACATGCGCTCTGCTCCCTTGTCTCCAGCTGACACAGC 302
QY 942 TCTGACCTTCTGACGCGAGCTGAGAAATGACCTGCGCCGCGAGCCCTACTCTGTGTA 1001
DB 301 TCTGACCTTCTGACGCGAGCTGAGAAATGACCTGCGCCGCGAGCCCTACTCTGTGTA 242
QY 1002 ATGAATTAAGGCGCTGCTGTGTCTGTG 1029
DB 241 ATGAATTAAGGCGCTGCTGTGTGTG 214
```

RESULT 13

```
US-10-154-884B-5582/C
; Sequence 5582, Application US/10154884B
; Publication No. US2004005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/154, 884B
; PRIOR FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5582
; LENGTH: 421
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-5582
```

```
Query Match 19.6%; Score 208; DB 16; Length 421;
Best Local Similarity 100.0%; Pred. No. 2.8e-49;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 822 CTGGAAGTGGCTCTGGGGTCTCTACATGCGCTCTGCTCCCTTGTCTCCAGCTGACACAGC 881
DB 421 CTGGAAGTGGCTCTGGGGTCTCTACATGCGCTCTGCTCCCTTGTCTCCAGCTGACACAGC 362
QY 882 CTGGAAGTGGCTCTGGGGTCTCTACATGCGCTCTGCTCCCTTGTCTCCAGCTGACACAGC 941
DB 361 CTGGAAGTGGCTCTGGGGTCTCTACATGCGCTCTGCTCCCTTGTCTCCAGCTGACACAGC 302
```

```
QY 942 TCTGACCTTCTGACGCGAGCTGAGAAATGACCTGCGCCGCGAGCCCTACTCTGTGTA 1001
DB 301 TCTGACCTTCTGACGCGAGCTGAGAAATGACCTGCGCCGCGAGCCCTACTCTGTGTA 242
QY 1002 ATGAATTAAGGCGCTGCTGTGTCTGTG 1029
DB 241 ATGAATTAAGGCGCTGCTGTGTGTG 214
```

RESULT 14

```
US-10-764-324-5582/C
; Sequence 5582, Application US/10764324
; Publication No. US20040175739A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/764,324
; PRIOR FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5582
; LENGTH: 421
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-764-324-5582
```

```
Query Match 19.6%; Score 208; DB 17; Length 421;
Best Local Similarity 100.0%; Pred. No. 2.8e-49;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 822 CTGGAAGTGGCTCTGGGGTCTCTACATGCGCTCTGCTCCCTTGTCTCCAGCTGACACAGC 881
DB 421 CTGGAAGTGGCTCTGGGGTCTCTACATGCGCTCTGCTCCCTTGTCTCCAGCTGACACAGC 362
QY 882 CTGGAAGTGGCTCTGGGGTCTCTACATGCGCTCTGCTCCCTTGTCTCCAGCTGACACAGC 941
DB 361 CTGGAAGTGGCTCTGGGGTCTCTACATGCGCTCTGCTCCCTTGTCTCCAGCTGACACAGC 302
QY 942 TCTGACCTTCTGACGCGAGCTGAGAAATGACCTGCGCCGCGAGCCCTACTCTGTGTA 1001
DB 301 TCTGACCTTCTGACGCGAGCTGAGAAATGACCTGCGCCGCGAGCCCTACTCTGTGTA 242
QY 1002 ATGAATTAAGGCGCTGCTGTGTCTGTG 1029
DB 241 ATGAATTAAGGCGCTGCTGTGTGTG 214
```


RESULT 15

US-10-242-535A-9083
; Sequence 9083, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9083
; LENGTH: 400
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (378)..(378)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-9083

Query Match 19.4%; Score 205.2; DB 16; Length 400;

Best Local Similarity 97.0%; Pred.No.1.8e-48; Mismatches 4; Indels 3; Gaps 2;

Matches 230; Conservative 0; Mismatches 4; Indels 3; Gaps 2;

Qy	613	GATGCGAGCCGGAGATATGTGATGTGTC	CCAGAACTGATCTGAGCGGCTAAGACT	672
Db	164	GATGCGAGCCGGAGATATGTGATGTGTC	CCAGAACTGATCTGAGCGGCTAAGACT	223
Qy	673	GAGCTGCGCCCTGATGTTCCCAAGAGCA	AGAAAGTGGAGAAAGGGGCTCCAGAT	732
Db	224	GAGCTGCGCCCTGATGTTCCCAAGAGCA	AGAAAGTGGAGAAAGGGGCTCCAGAT	283
Qy	733	TACGAGATCTGCGAGAGCTGAATGAGGG	CGCTGTGAGGGCCGAGTCTCTCGAACA	792
Db	284	TACGAGATCTGCGAGAGCTGAATGAGGG	CGCTGTGAGGGCCGAGTCTCTCGAACA	343
Qy	793	GGCTTGCTGGGAC-GGCTGAGCTGGGCA	GTGAA--GTGGCTGTGGGCTCTCAC	846
Db	344	GGCTTGCTGGGAC-GGCTGAGCTGGGCA	GTGAA--GTGGCTGTGGGCTCTCAC	400

Search completed: November 22, 2004, 13:40:22
Job time : 617 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 07:46:40 ; Search time 6715 Seconds
(without alignments)
5752.214 Million cell updates/sec

Title: US-09-597-920B-1

Perfect score: 1060

Sequence: 1 gacctgccttgaggggcgc.....aaaaaaaaaaaaaaaaaaaa 1060

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: gb_ests1.*
2: gb_ests2.*
3: gb_hnc.*
4: gb_ests3.*
5: gb_ests4.*
6: gb_ests5.*
7: gb_ests6.*
8: gb_gsa1.*
9: gb_gsa2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1027.4	96.9	1534	3	CR624409 full-length
2	1027.4	96.9	1564	3	CR597721 full-length
3	1027.4	96.9	1604	3	CR592470 full-length
4	968.4	91.4	1207	3	CR602091 full-length
5	814.8	76.9	1090	4	BM810074 AGENCOURT
6	666	62.8	702	9	AY403398 Homo sapi
7	645.2	60.9	896	1	AL558417
8	633.6	58.8	983	1	AL558417
9	632.4	59.7	937	4	BG751757
10	625.6	59.0	935	1	AL580592
11	625.2	59.0	922	1	AL580240
12	612.2	57.8	651	2	AM612429
13	612	57.7	878	1	AL558866
14	606.4	57.2	610	4	BM715456
15	578	54.5	811	5	BM437371
16	575.6	54.3	916	5	BM447994
17	575	54.2	775	5	BI753306
18	572	54.0	846	5	BM437370
19	549	51.8	771	4	BI838499
20	547	51.6	924	4	BI910865
21	543.2	51.2	1096	4	BM562211
22	534.4	50.4	544	6	CD370683
23	525.6	49.6	1080	5	BQ072651
24	524.4	49.5	884	5	BM369483

25	519.4	49.0	817	4	BI911404
26	519.2	49.0	702	9	AY403399
27	516.8	48.8	874	4	BI909045
28	512	48.3	630	2	BF906497
29	509	48.0	517	4	BM673935
30	493.6	46.6	929	4	BM386066
31	477.6	45.1	540	5	BO581351
32	477.2	45.0	826	4	BI838054
33	467.4	44.1	643	2	AM384820
34	466	44.0	665	2	AM384832
35	459.8	43.4	568	2	AM970405
36	454	42.8	518	1	AI922013
37	445	42.0	898	5	BM400523
38	444.4	41.9	816	4	BI911769
39	440	41.5	456	1	AM002632
40	433.2	40.9	649	6	CA426247
41	432.2	40.8	477	2	AM296085
42	429.6	40.5	450	2	BF508034
43	423.8	40.0	657	4	BI909736
44	420.8	39.7	827	4	BM397816
45	418	39.4	892	5	BM330901

ALIGNMENTS

RESULT 1	CR624409	1534 bp	mRNA	linear	HTC 21-JUL-2004
LOCUS	CR624409				
DEFINITION	full-length cDNA clone CSOCAP06YF09 of Thymus of Homo sapiens (human)				
ACCESSION	CR624409.1	GI:50505216			
VERSION	CR624409.1				
KEYWORDS	HTC; CNSLT; cDNA.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
AUTHORS	Li, W.B., Gruber, C., Jessee, J., and Polayes, D.				
JOURNAL	Full-length cDNA libraries and normalization				
REMARK	Unpublished				
	Contact: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paraday Avenue				
	2 (bases 1 to 1534)				
REFERENCE	Genoscope.				
AUTHORS	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage ;				
TITLE	BP 191 9106 EVR cedex - FRANCE (E-mail: sequef@genoscope.cns.fr				
JOURNAL	- Web : www.genoscope.cns.fr				
COMMENT	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.				
FEATURES	Location/Qualifiers				
source	1..1534				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="CSOCAP06YF09"				
	/tissue_type="Thymus"				
	/plasmid="pCMVSPORT_6"				
ORIGIN					
Query Match	96.9%; Score 1027.4; DB 3; Length 1534;				
Best Local Similarity	99.9%; Fred. No. 4.2e-229;				
Matches 1028; Conservative	0; Mismatches 1; Indels 0; Gaps 0;				
1	GACCTGCTTGAAGGGGCTAGAGGCTGACCCGAGCTCCGAGCTCCCTGAGATG 60				
234	GACCTGCTTGAAGGGGCTAGAGGCTGACCCGAGCTCCGAGCTCCCTGAGATG 293				

QY	61	GAGGAGGGCCATCTGTGTCCTCCCTGGCGGTGGGGGCTCTGTGTCGTGTCGCATCTGGCCATG	120
Db	294	GAGGAGGGCCATCTGTGTCCTCCCTGGCGGTGGGGGCTCTGTGTCGTGTCGCATCTGGCCATG	353
QY	121	TTGATGGGACGTGTGTGTGTGACCTGGACACAGACTGGCCAGGGCTCTTAAGAGACATCTCTCA	180
Db	354	TTGATGGGACGTGTGTGTGTGACCTGGACACAGACTGGCCAGGGCTCTTAAGAGACATCTCTCA	413
QY	181	GATAGTTTGTATCCAAAGGGGCAATCCAGTTCAAAAGGCTCACAGGGTTGCCCTTGCGCA	240
Db	414	GATAGTTTGTATCCAAAGGGGCAATCCAGTTCAAAAGGCTCACAGGGTTGCCCTTGCGCA	473
QY	241	CCTGCTTACCCACGTGACCTCTACCCACCCCTGAGCGAGCCAGACTGCTGCCCATC	300
Db	474	CCTGCTTACCCACGTGACCTCTACCCACCCCTGAGCGAGCCAGACTGCTGCCCATC	533
QY	301	CCAGATCCCGCAGCCCTTGGGGGCTTCCACCGGACGCCATCTTCCCGCGGAGTTCT	360
Db	534	CCAGATCCCGCAGCCCTTGGGGGCTTCCACCGGACGCCATCTTCCCGCGGAGTTCT	593
QY	361	GATGTGCCAACAAGTGTGGCGAGCTACAGAAACAGAAACAGAGCTGTGAGATCCAGAT	420
Db	594	GATGTGTCCCAAGAAGTGTGGCGAGCTACAGAAACAGAAACAGAGCTGTGAGATCCAGAT	653
QY	421	GAGATGAGAGACGACTATCAACAACCGAGCTACGTGAGTGTCTCTGACAGCACCCCG	480
Db	654	GAGATGAGAGACGACTATCAACAACCGAGCTACGTGAGTGTCTCTGACAGCACCCCG	713
QY	481	GCCACTAGACCTGTGCTGCCCATACGCTCTCTGACTACAGACCCCTTGACATCCGACAGT	540
Db	714	GCCACTAGACCTGTGCTGCCCATACGCTCTCTGACTACAGACCCCTTGACATCCGACAGT	773
QY	541	GCGTCTCCATGAGATCCATTTGATATGACGAAAGTTCCGGAAGAGGGGGAGAGCGCA	600
Db	774	GCGTCTCCATGAGATCCATTTGATATGACGAAAGTTCCGGAAGAGGGGGAGAGCGCA	833
QY	601	GAAGCGTCTCTGAGATGGCAGCCGGGAGTATGTGAATGTGTCCCAAGAACTGCATCTTGA	660
Db	834	GAAGCGTCTCTGAGATGGCAGCCGGGAGTATGTGAATGTGTCCCAAGAACTGCATCTTGA	893
QY	661	GCGGCTTAAGACTGAGCTGTGCCGCCCTGAGTTCCCAAGAGGACAGAGAACTGAGAGAG	720
Db	894	GCGGCTTAAGACTGAGCTGTGCCGCCCTGAGTTCCCAAGAGGACAGAGAACTGAGAGAG	953
QY	721	GGGGCTCCAGATTAGAGAAATCTGAGAGAGCTGAAGGGCTGTGAGGGCCGAGTCT	780
Db	954	GGGGCTCCAGATTAGAGAAATCTGAGAGAGCTGAAGGGCTGTGAGGGCCGAGTCT	1013
QY	781	GTCCTGGAACCAAGGCTTGGCTTGGGAGCGGCTGAGCTGGGAGCTGGAATGTGCTTGGGGT	840
Db	1014	GTCCTGGAACCAAGGCTTGGCTTGGGAGCGGCTGAGCTGGGAGCTGGAATGTGCTTGGGGT	1073
QY	841	CCTCAACATGGCGTCTGTGCCCTTGTCTCAAGCTTGACACACAGCTTGAGAAATCCCCCGTAA	900
Db	1074	CCTCAACATGGCGTCTGTGCCCTTGTCTCAAGCTTGACACACAGCTTGAGAAATCCCCCGTAA	1133
QY	901	CTTATTATCACTTTGAGGGTTGGGCTGTGTGCCCGGAAGGCTGTGCACTTCTGACGAG	960
Db	1134	CTTATTATCACTTTGAGGGTTGGGCTGTGTGCCCGGAAGGCTGTGCACTTCTGACGAG	1199
QY	961	CTTGAAGATGACCTTGCCTCCGTGCCCGCAGCCCTTACTGTGTATATAGAAATAAAGGCTTGGCT	1020
Db	1194	CTTGAAGATGACCTTGCCTCCGTGCCCGCAGCCCTTACTGTGTATATAGAAATAAAGGCTTGGCT	1255
QY	1021	GTGTCTGTG 1029	
Db	1254	GTGTCTGTG 1262	

DEFINITION	full-length cDNA clone CSDDM005Yp13 of Fetal liver of Homo sapiens (human).
ACCESSION	CR597721
VERSION	CR597721.1 GI:50478528
KEYWORDS	HTC; cNSLT cDNA.
SOURCE	
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE	1 (bases 1 to 1564)
JOURNAL	Li, W.B., Gruber, C., Jessee, J. and Polayars, D.
REMARK	Full-length cDNA libraries and normalization Unpublished Contract : Peng Liang Email : fliang@lifeepoch.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue 2 (bases 1 to 1564)
REFERENCE	Genoscope.
AUTHORS	Direct Submission
TITLE	Submitted (20-JUN-2004) Genoscope - Centre National de Sequencage
JOURNAL	Bp 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT	1st strand cDNA was primed with a NotI-oligo (dt) primer. Five priming end enriched, double-strand cDNA was digested with Not I and clone into the Not I and Ecor V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. Location/Qualifiers 1. 1564 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CSDDM005Yp13" /tissue_type="Fetal liver" /plasmid="pCMVSPORT_6"
FEATURES	
source	1. 1564
ORIGIN	
Query Match	96.9% Score 1027.4; DB 3; Length 1564;
Best Local Similarity	99.9% Pred. No. 4.2e-229;
Matches 1028; Conservative	0; Mismatches 1; Indels 0; Gaps 0
1	GACTCTGCCTTTGAGGGGCTTGAAGGGTGCAGGCAGCCTTCTCGAAGTCCCTGCGAGATG 60
238	GACTCTGCCTTTGAGGGGCTTGAAGGGTGCAGGCAGCCTTCTCGAAGTCCCTGCGAGATG 297
61	GAGGAGGCCATCTCTGCTCCCTGCTGCTGCTGGGGCTTCTGCTGCTGCCATCTTGCCATG 120
238	GAGGAGGCCATCTCTGCTCCCTGCTGCTGCTGGGGCTTCTGCTGCTGCCATCTTGCCATG 357
121	TTGATGGGACGTGTGTGTGACATCTGCACACAGCTGCCAGGCTCTTACGACAGCATCTCTCA 180
358	TTGATGGACATGTGTGTGTGACATCTGCACACAGCTGCCAGGCTCTTACGACAGCATCTCTCA 417
181	GATAGTTTGTATCCAAAGGGGACATCCAGTTCAACGCGCTCAACGCGTTGCCCTTGCCCA 240
418	GATAGTTTGTATCCAAAGGGGACATCCAGTTCAACGCGCTCAACGCGTTGCCCTTGCCCA 477
241	CCTGCTTACCACTCTGTACATCTTCTTACCAACCCCTGAGCGAGCAGACTGCTCCCCATG 300
478	CCTGCTTACCACTCTGTACATCTTCTTACCAACCCCTGAGCGAGCAGACTGCTCCCCATG 537
301	CCAAAGTCCCGGACAGCCCTTGGGGGGCTCCCAACGCGAGCGGCATCTTCCCGGGGGATTCT 360
538	CCAAAGTCCCGGACAGCCCTTGGGGGGCTCCCAACGCGAGCGGCATCTTCCCGGGGGATTCT 597
361	GATGTGCTCAACAGTGTGGCGAGCTACGAGAAACGAGAAACAGCGCTGTGAGATCAAGT 420
598	GATGTGCTCAACAGTGTGGCGAGCTACGAGAAACGAGAAACAGCGCTGTGAGATCAAGT 457
421	GAGATGAGAGACGACTATTCACAAACCAAGCTTACCTGTGTGGCTTCTGACAGCAACCCCG 480
658	GAGATGAGAGACGACTATTCACAAACCAAGCTTACCTGTGTGGCTTCTGACAGCAACCCCG 717

QY 481 GCCACTAGACCTGCTGCCCCATCAGCTCTGCACTCAGACACCCCTGGCATCCGACAGT 540
 |||||
 DB 718 GCCACTAGACCTGCTGCCCCATCAGCTCTGCACTCAGACACCCCTGGCATCCGACAGT 777
 QY 541 GCGTTCTCATGAGTCCATTGATGATTAAGTGAACGTTTCCGAGAGCGGGAGAGCGCA 600
 |||||
 DB 778 GCGTTCTCATGAGTCCATTGATGATTAAGTGAACGTTTCCGAGAGCGGGAGAGCGCA 837
 QY 601 GAAGGCTCTGTGATGAGGAGCGGGAGTATGTAATGTGTCCAGAACTGCTCTCTGA 660
 |||||
 DB 838 GAAGGCTCTGTGATGAGGAGCGGGAGTATGTAATGTGTCCAGAACTGCTCTCTGA 897
 QY 661 GCGGCTTAAGTGAAGCTGCGCCCTGAGTTCACAGAGGAGAGAGTGAAGAGAG 720
 |||||
 DB 898 GCGGCTTAAGTGAAGCTGCGCCCTGAGTTCACAGAGGAGAGAGTGAAGAGAG 957
 QY 721 GCGGCTTAAGTGAAGTGAAGTCTGAGAGAGTGAAGGCTGTGAGAGCGGAGTCT 780
 |||||
 DB 958 GCGGCTTAAGTGAAGTGAAGTCTGAGAGAGTGAAGGCTGTGAGAGCGGAGTCT 1017
 QY 781 GTCCTGGAACAGAGCTTGTGCTGGAGAGGCTGAGCTGGAGAGTGTCTTGGGGT 840
 |||||
 DB 1018 GTCCTGGAACAGAGCTTGTGCTGGAGAGGCTGAGCTGGAGAGTGTCTTGGGGT 1077
 QY 841 CCTCAATGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 |||||
 DB 1078 CCTCAATGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1137
 QY 901 CTATATATCATCTTTGGGGTTCGCGCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 |||||
 DB 1138 CTATATATCATCTTTGGGGTTCGCGCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1197
 QY 961 CCTGGAATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 |||||
 DB 1198 CCTGGAATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1257
 QY 1021 GTGTCTGTG 1029
 |||||
 DB 1258 GTGTCTGTG 1266

RESULT 3
 CRS92470 1604 bp mRNA linear HTC 21-JUL-2004
 LOCUS
 DEFINITION Full-length cDNA clone CS0D015YE16 of T cells (Jurkat cell line)
 Cot 10-normalized of Homo sapiens (human).
 CR592470
 ACCESSION
 VERSION CRS92470.1 GI:50473277
 KEYWORDS HTC; CNSLT cDNA.
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 1604)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished
 Contact: Feng Liang Email: fliang@life.techn.com URL:
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue
 2 (bases 1 to 1604)
 Genoscope.
 Direct Submission
 Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail: sequef@genoscope.cns.fr
 - Web: www.genoscope.cns.fr)
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 Location/Qualifiers
 1..1604

ORIGIN
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D015YE16"
 /issue_type="T cells (Jurkat cell line) Cot
 10-normalized"
 /plasmid="pCMVSPORT_6"

Query Match 96.9%; Score 1027.4; DB 3; Length 1604;
 Best Local Similarity 99.9%; Pred. No. 4.2e-229;
 Matches 1028; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACTGCTGCTTGAAGGAGGCTTGAAGGAGGCTTGAAGGAGGCTTGAAGGAGGCTTGAAGGAGG 60
 |||||
 DB 240 GACTGCTGCTTGAAGGAGGCTTGAAGGAGGCTTGAAGGAGGCTTGAAGGAGGCTTGAAGGAGG 299
 QY 61 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
 |||||
 DB 300 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 359
 QY 121 TTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
 |||||
 DB 360 TTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 419
 QY 181 GATAGTTGTATCCAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
 |||||
 DB 420 GATAGTTGTATCCAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 479
 QY 241 CCTGCTTACCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 |||||
 DB 480 CCTGCTTACCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
 QY 301 CCAAGATCCCGCAGACCCCTTGGGGGCTTCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
 |||||
 DB 540 CCAAGATCCCGCAGACCCCTTGGGGGCTTCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 599
 QY 361 GATGTTGCCAAGTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
 |||||
 DB 600 GATGTTGCCAAGTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 659
 QY 421 GAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
 |||||
 DB 660 GAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 719
 QY 481 GCCACTAGACCTGCTGCCCCATCAGCTCTGCACTCAGACACCCCTGGCATCCGACAGT 540
 |||||
 DB 720 GCCACTAGACCTGCTGCCCCATCAGCTCTGCACTCAGACACCCCTGGCATCCGACAGT 779
 QY 541 GCGTTCTCATGAGTCCATTGATGATTAAGTGAACGTTTCCGAGAGCGGGAGAGCGCA 600
 |||||
 DB 780 GCGTTCTCATGAGTCCATTGATGATTAAGTGAACGTTTCCGAGAGCGGGAGAGCGCA 839
 QY 601 GAAGGCTCTGTGATGAGGAGCGGGAGTATGTAATGTGTCCAGAACTGCTCTCTGA 660
 |||||
 DB 840 GAAGGCTCTGTGATGAGGAGCGGGAGTATGTAATGTGTCCAGAACTGCTCTCTGA 899
 QY 661 GCGGCTTAAGTGAAGCTGCGCCCTGAGTTCACAGAGGAGAGAGTGAAGAGAG 720
 |||||
 DB 900 GCGGCTTAAGTGAAGCTGCGCCCTGAGTTCACAGAGGAGAGAGTGAAGAGAG 959
 QY 721 GCGGCTTAAGTGAAGTGAAGTCTGAGAGAGTGAAGGCTGTGAGAGCGGAGTCT 780
 |||||
 DB 960 GCGGCTTAAGTGAAGTGAAGTCTGAGAGAGTGAAGGCTGTGAGAGCGGAGTCT 1019
 QY 781 GTCCTGGAACAGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 |||||
 DB 1020 GTCCTGGAACAGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079
 QY 841 CCTCAATGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 |||||
 DB 1080 CCTCAATGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1139

QY 901 CTTATTATCACTTTGGGGTTGGCGCTGTGTCCCGGAAAGCTTGACACTTCTGACGAG 960
DB 1140 CTTATTATCACTTTGGGGTTGGCGCTGTGTCCCGGAAAGCTTGACACTTCTGACGAG 1199
QY 961 CTTGAGAAATGACCTTGGCGCTTGGCGCGCAAGCTTCTGTGTATATGAATAAGGCTTGGCT 1020
DB 1200 CTTGAGAAATGACCTTGGCGCTTGGCGCGCAAGCTTCTGTGTATATGAATAAGGCTTGGCT 1259
QY 1021 GTGTCTGTG 1029
DB 1260 GTGTCTGTG 1268

RESULT 4
CR602091 1207 bp mRNA linear HTC 21-JUL-2004
LOCUS DEFINITION full-length cDNA clone CS0DU05YC09 of T cells (Jurkat cell line)
CR602091 COT 10-normalized of Homo sapiens (human).
ACCESSION CR602091.1 GI:50482898
VERSION HTC; CNS1T.CDNA
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1. (bases 1 to 1207)
AUTHORS Li, W. B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.lifetech.com/Invitrogen Corporation 1600
Paradise Avenue
2. (bases 1 to 1207)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequence :
JOURNAL BP 191 91006 Evry cedex - FRANCE (E-mail: seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(AT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source location/Qualifiers
1. .1207
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DU05YC09"
/tissue_type="T cells (Jurkat cell line) Cot
10-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 91.4%; Score 968.4; DB 3; Length 1207;
Best Local Similarity 99.9%; Pred. No. 2.4e-215;
Matches 969; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 241 CTGCTACCCACCTGTGCACTCTCTACCCACCCCTGAGCAGCAGACTGTCCCATC 300
DB 478 CTGCTACCCACCTGTGCACTCTCTACCCACCCCTGAGCAGCAGACTGTCCCATC 537
QY 301 CCAAGATCCCGGAGCCCTTGGGGGCTCCACCGAGCGCATCTTCCGGGGATTC 360
DB 538 CCAAGATCCCGGAGCCCTTGGGGGCTCCACCGAGCGCATCTTCCGGGGATTC 597
QY 361 GATGATGCCCAAGCTGTGGGAGCTACGAAAGAGGAACACCTGTGAGATGCAT 420
DB 598 GATGATGCCCAAGCTGTGGGAGCTACGAAAGAGGAACACCTGTGAGATGCAT 657
QY 421 GAGGATGAGACGACTATCAACACCCAGGCTAGCTGTGCTTCCCTGACAGACCCG 480
DB 658 GAGGATGAGACGACTATCAACACCCAGGCTAGCTGTGCTTCCCTGACAGACCCG 717
QY 481 GCCACTAGCACTGTGCGCCCATGACTCTGTGCACTGACACCCCTGATCCAGACAT 540
DB 718 GCCACTAGCACTGTGCGCCCATGACTCTGTGCACTGACACCCCTGATCCAGACAT 777
QY 541 GCTTCTCCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 778 GCTTCTCCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 837
QY 601 GAAGCGTCTGTGATGAGCAGCCGAGAGTATGATGATGATGATGATGATGATGATGAT 660
DB 838 GAAGCGTCTGTGATGAGCAGCCGAGAGTATGATGATGATGATGATGATGATGATGAT 897
QY 661 GCGGCTTAAGACTGAGCTGCGCCCTGAGTTCCAGAGGCGAGAGAGTGAAGAG 720
DB 898 GCGGCTTAAGACTGAGCTGCGCCCTGAGTTCCAGAGGCGAGAGTGAAGAG 957
QY 721 GGGGCTTCAATTAACGAGATCTGCAAGAGCTGAACCTGAGGGCTGTGAGGCGAGTCT 780
DB 958 GGGGCTTCAATTAACGAGATCTGCAAGAGCTGAACCTGAGGGCTGTGAGGCGAGTCT 1017
QY 781 GTCTGGAAACGAGCTTGCCTGAGGAGCGCTGAGCTGAGCTGAGCTGAGCTGAGCT 840
DB 1018 GTCTGGAAACGAGCTTGCCTGAGGAGCGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1077
QY 841 CTTCAATGCGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
DB 1078 CTTCAATGCGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1137
QY 901 CTTATTATCACTTTGGGGTTGGCGCTGTGTCCCGGAAAGCTTGACACTTCTGACGAG 960
DB 1138 CTTATTATCACTTTGGGGTTGGCGCTGTGTCCCGGAAAGCTTGACACTTCTGACGAG 1197
QY 961 CTTGAGAAATG 970
DB 1198 CTTGAGAAATG 1207

RESULT 5
BM810074 1090 bp mRNA linear EST 05-MAR-2002
LOCUS AGENCOURT_6579672 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5470018
DEFINITION 5', mRNA sequence.
ACCESSION BM810074
VERSION BM810074.1 GI:19126897
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1. (bases 1 to 1090)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@ds-r@mail.nih.gov
Tissue Procurement: DCTD/DTF

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

FEATURES

SOURCE

```

1.. 1090
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5470018"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="VDH10 (phage-resistant)"
/clone_11b="NH MGC_41"
/note="Organ: skin; Vector: pOT87; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NH MGC Library."

```

ORIGIN

Query Match	76.9%;	Score 814.8;	DB 4;	Length 1090;
Best Local Similarity	94.1%;	Pred. No. 1.9e-179;		
Matches 851; Conservative	0;	Mismatches 52;	Indels 1;	Gaps 1

Qy	1	GGGGCCTAAGGGAGTGCAGCAGCCGCTCCGACACTCCCTGCAGATGAGAGAGGCCATCCT	74
Db	76	GGGGCCTTAAGGGATGCAGCCAGCCTGCTCCGAACTCCCTGCAGATGAGAGAGGCCATCCT	135
Qy	75	GGTCCCTGCTGCTGCTGGGGCTCTCTGCTGCTGCCATCTTGCCATGTTGATGAGCACTGTG	134
Db	136	GGTCCCTGCTGCTGCTGGGGCTCTCTGCTGCTGCCATCTTGCCATGTTGATGAGCACTGTG	195
Qy	135	TGTGCACCTGCACAGACTCCAGGGCTCCCTACAGACAGCAATCTCCAGATATTTGTTATCC	194
Db	196	TGTGCACCTGCACAGACTCCAGGGCTCTCAGACAGCAATCTCAGATATTTGTTATCC	255
Qy	195	AAGGGGCATTCAGTTTCAAACGGGCTTCACACGGTTGCCCTTGCGCCACTGCTTACCCAGC	254
Db	256	AAGGGGCATTCAGTTTCAAACGGGCTTCACACGGTTGCCCTTGCGCCACTGCTTACCCAGC	315
Qy	255	TGTACACTCTCTAACCCACCCCTTAGACGACAGCAACTGCTCTCCCATCCCAATATCCCGCA	314
Db	316	TGTACACTCTCTAACCCACCCCTTAGACGACAGCAACTGCTCTCCCATCCCAATATCCCGCA	375
Qy	315	GGCCCTTGGGGGGCTCCACACGGACGCATCTTCCGGGGGAGTTCTGATGTGTCCAAACAG	374
Db	376	GGCCCTTGGGGGGCTCCACACGGACGCATCTTCCGGGGGAGTTCTGATGTGTCCAAACAG	435
Qy	375	TGTGGGAGCTTACAGAAACAGAGAAACAGACTGTGTAGATGTCAATGATGAGATGAGACGA	434
Db	436	TGTGGGAGCTTACAGAAACAGAGAAACAGACTGTGTAGATGTCAATGATGAGATGAGACGA	495
Qy	435	CTATCAACAACCAAGGCTACTCGTGGTGTCTTCTGACAGACCCCGGCACTTACGACTGC	494
Db	496	CTATCAACAACCAAGGCTACTCGTGGTGTCTTCTGACAGACCCCGGCACTTACGACTGC	555
Qy	495	TGCCCCATCAGCTCTGCACTGACACCCCTTGAGATCCGAGACAGTGCCTTCTCCATATGA	554
Db	556	TGCCCCATCAGCTCTGCACTGACACCCCTTGAGATCCGAGACAGTGCCTTCTCCATATGA	615
Qy	555	GTCATTATGATTAAGTGAACGTTTCCGGAGAGCGGGAGAGCGAGAAAGGTCCTGTGA	614
Db	616	GTCATTATGATTAAGTGAACGTTTCCGGAGAGCGGGAGAGCGAGAAAGGTCCTGTGA	675
Qy	615	TGGCAGCCGGAGATATGTGAATGTGTCCAGGACTGACATCTGTGAGCGGCTTAAGACTGA	674

Db	676	TGGACAGCCGGAGTAGTAGTAATGTGTCCAGAAATCGATCCTG9AGGGGCTAAGACTGA	725
OY	675	GCCCTGCCCCCTTGAGTTCCAGAGGACAGAGAAATGAGAGAAAGGGGGCTCCAGATT	724
Db	736	GCCCTGCCCCCTTGAGTTCCAGAGGACAGAGAAATGAGAGAAAGGGGGCTCCAGATT	724
OY	735	CGAAGATCTGCAGAGAGCTGAATCTGAGAGGCTGTGGAGGCGCAGTCTGTCTGAAACAGG	724
Db	795	CNAGATCTGCAGAGAGCTGAATCTGAGAGGCTGTGGAGGCGCAGTCTGTCTGAAACAGG	854
OY	795	CTTGCCCTGGGACCGGCTGAGCTGGGCACTGGAAATGAGCTTCTGGGGTCTCAATGGCGTC	854
Db	855	CTTGCCCTGGGACCGGCTGAGCTGGGCACTGGAAATGAGCTTCTGGGGTCTCAATGGCGTC	914
OY	855	CTGACCCTTGCTCCAGCTGACAAACAGCCTGGAAGAAATCCCCCGTAACTTAATTACATTT	914
Db	915	CTGGCCCTTGCTCCAGCTGACAAACAGCCTGGAAGAAATCCCCCGTAACTTAATTACATTT	974
OY	915	GGGG 918	
Db	975	TGGG 978	

RESULT 6	AY403398	702 bp	DNA	linear	GSS 15-DEC-2001
LOCUS	AY403398				
DEFINITION	Homo sapiens LAT gene, VIRTUAL TRANSCRIPT, partial sequence,				
ACCESSION	AY403398				
VERSION	AY403398.1	GI:39759381			

SOURCE ORGANISM	Homo sapiens (human)	Homo sapiens
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

REFERENCE
AUTHORS

Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 702)
Clark, A. G., Głanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,

TITLE Inferring nonneutral evolution from gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)

TITLE	JOURNAL	COMMENT
Direct Submission	Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA	Submitted (16-NOV-2003) This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES	Location/Qualifiers
source	1. .702

ORIGIN

Query	58	ATGAGAGAGGCACATCCCTGCTCCCTGCGCTGCGGGGCTCTGCTGCTGCTGCCATCTCTGGCC	117
Db	1	ATGAGAGAGGCACATCCCTGCTCCCTGCGCTGCTGAGGGCTCTGCTGCTGCTGCCATCTCTGGCC	60
Qy	118	ATGTTGATGTCACATGTTGTGTCACATGTCACACAGATCTGCCAGGCTCTCTAGCAGCAGCATCC	177

[illegible]

RESULT 7	LOCUS	DEFINITION	ACCESSION	VERSION	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
AL558417	AL558417	996 bp mRNA linear	AL558417	AL558417	AL558417.3	GI:46183815	EST.	Homo sapiens (human)	Homo sapiens	Eumalvota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	1. (bases 1 to 996) I. W.B., Gruber, C., Jesssee, J. and Polayes, D. Full-length cDNA libraries and normalization Unpublished (2001) On Feb 15, 2001 this sequence version replaced gi:31280216.
											<p>Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr</p> <p>1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primer end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 8019.f</p> <p>For more information about this cluster, see http://www.genoscope.cns.fr/cdna?s=CS05D0055B05SQP1&c=8019.f.</p>

[illegible]

SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Homo sapiens (human)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	1	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.	Full-length cDNA libraries and normalization	Unpublished (2001)	On May 22, 2003 this sequence version replaced gi:31027882.
Genoscope	Contact: Genoscope	Genoscope - Centre National de Sequencage	BP 191 91006 EVRY cedex - France	Email: sequef@genoscope.cns.fr. Web : www.genoscope.cns.fr	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.	This sequence belongs to sequence cluster 8019.f
For more information about this cluster, see	http://www.genoscope.cns.fr/cdna?b=CS0AM005CH07QPI&c=8019.f.	Location/Qualifiers	1..983	/organism="Homo sapiens"	/mol_type="mRNA"	/db_xref="taxon:9606"
/clone="CS0DM005IP13"	/tissue_type="FETAL LIVER"	/dev_stage="fetal"	/clone.lib="Homo sapiens FETAL LIVER"	/note="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."		
ORIGIN						
Query Match	59.8%;	Score 633.6;	DB 5;	Length 983;		
Best Local Similarity	98.6%;	Pred. No. 3.9e-137;				
Matches	655;	Conservative	4;	Mismatches	3;	Indels
					2;	Gaps
1	GACTCTGCCCTTGTAGAGGGGCTTGTAGAGGGTGTAGAGGCTGTCCGAGCTCCCTGTGAGATG	60				
239	GACTCTGCCCTTGTAGAGGGGCTTGTAGAGGGTGTAGAGGCTGTCCGAGCTCCCTGTGAGATG	298				
61	GAGGAGGCGCATCTCTGTCTCCCTGCTGCTGGGGCTCTCTGCTGCTGCCATCTTGCCCATG	120				
299	GAGGAGGCGCATCTCTGTCTCCCTGCTGCTGGGGCTCTCTGCTGCTGCCATCTTGCCCATG	358				
121	TTGATGGCACTGTGTGTGTCACATGTCACAGATCGGCGGCTCTCTAAGACAGATATCTCTCA	180				
359	TTGATGGCACTGTGTGTGTCACATGTCACAGATCGGCGGCTCTCTAAGACAGATATCTCTCA	418				
181	GATAGTTTGTATCCAAAGGGGATCCAGTTTCAAACGGGCTCTCAACAGGTTTCCCTGTGGCA	240				
419	GATAGTTTGTATCCAAAGGGGATCCAGTTTCAAACGGGCTCTCAACAGGTTTCCCTGTGGCA	478				
241	CTGTGCTTACCACTGTGTACCTCTCTTCAACCCCTGTAGCCAGCCAGATCTGCTCCCATC	300				
479	CTGTGCTTACCACTGTGTACCTCTCTTCAACCCCTGTAGCCAGCCAGATCTGCTCCCATC	538				
301	CCAAGATCCCGCAGAGCCCTTGTGGGGGCTCCAGCCGGAAGCATCTTCCCGGCGGGAATCT	360				
539	CCAAGATCCCGCAGAGCCCTTGTGGGGGCTCCAGCCGGAAGCATCTTCCCGGCGGGAATCT	598				
361	GATGTGTCCCAACAGTGTGGGAGCTACGAGAACGAGAACCAAGCCTGTGAGATGACAT	420				
599	GATGTGTCCCAACAGTGTGGGAGCTACGAGAACGAGAACCAAGCCTGTGAGATGACAT	658				
421	GAGGATGAGGACGACTATCAACAACCCAGGCTTACCTGTGTGTGCTTCTCGAAGACACCCG	480				
659	GAGGATGAGGACGACTATCAACAACCCAGGCTTACCTGTGTGTGCTTCTCGAAGACACCCG	718				

Qy		481	GCGACTAGACATGCTGGCCCCCATCAGCTCTGTGCATCAGACACCCTTGACATCCGAGCACT	540
Db		719	GCCATTAGACACTGCTGCCCCCATCACCTCTGTGCATCAGACACCCTTGACATCCGAGCACT	778
Qy		541	GCCCTTCATCGAGTATCCATTTGATGATTAACGGAACGTTCCGGAGAAGCGGGAGAGCGCA	600
Db		779	CCCTTCTCATAGAGTCCATTTGATGATTAACGGAACGTTCCGGAGAAGCGGGAGAGCGSC--	-837
Qy		601	GAAAGCTCTGTGCATGATGAGACGCCGAGATATGTAATGTGTCCCAAGAACTGCATCTTGA	660
Db		838	GAAAGGTCTCTGATGAGGACGCCGAGATTTGTTATGTGT-CCAGAACTGCATCTTGA	896
Qy		661	GGCG 664	
Db		897	GGCG 900	
RESULT 9				
Bg751757				
LOCUS				
DEFINITION		Bg751757	635 bp mRNA linear EST 15-MAY-2001	
ACCESSION		602730483F1 NIH_MGC_43	Homo sapiens cDNA clone IMAGE:4874108 5',	
VERSION		Bg751757	mRNA sequence.	
KEYWORDS		Bg751757	GI:14062410	
SOURCE		EST.		
ORGANISM		Homo sapiens (human)		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE		1 (bases 1 to 635)		
JOURNAL		NIH-MGC http://mgc.nci.nih.gov/.		
COMMENT		National Institutes of Health, Mammalian Gene Collection (MGC)		
		Unpublished (1999)		
		Contact: Robert Strausberg, Ph.D.		
		Email: cgabbs@omail.nih.gov		
		Tissue Procurement: ATCC		
		cDNA Library Preparation: Ling Hong/Rubin Laboratory		
		CNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)		
		DNA Sequencing by: Incyte Genomics, Inc.		
		Clone distribution: MGC clone distribution information can be		
		found through the I.M.A.G.E. Consortium/LIML at:		
		http://image.liml.gov		
		Plate: LILCM1753 row: e column: 21		
		High quality sequence stop: 635.		
FEATURES		Location/Qualifiers		
Source		1..635		
		/organism="Homo sapiens"		
		/mol_type="mRNA"		
		/db_xref="taxon:9606"		
		/clone="IMAGE:4874108"		
		/tissue_type="normal pigmented retinal epithelium"		
		/lab_host="DH10B (phage-resistant)"		
		/clone_id="NIH_MGC_43"		
		/note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2:		
		EcotRI; cDNA made by oligo-dT priming. Directionally		
		cloned into EcotRI/XhoI sites using the following 5'		
		adaptor: GGCACGAC(G) library constructed by Ling Hong		
		in the laboratory of Gerald W. Rubin (University of		
		California, Berkeley) using ZAP-cDNA synthesis kit		
		(Stratagene) and Superscript II RT (Life Technologies).		
		Note: this is a NIH_MGC Library. "		
ORIGIN				
Query Match		59.7%; Score 632.4; DB 4; Length 635;		
Best Local Similarity		99.8%; Pred. No. 7.1e-137;		
Matches 633; Conservative 1; Mismatch 1; Indels 0; Gaps 0;				
Qy		126	GGCACTGTGTGTGCACTGCCACAGACATGCCAGGCTCTTAGACAGACATCTCAGATAG	185
Db		2	GGCACTGTGTGTGCACTGCCACAGACATGCCAGGCTCTTAGACAGACATCTCAGATAG	61
Qy		186	TTTGTATCAAGGGGACATCCAGTTCAACGGCTCACACGGTTGCCCTTGACCACTGC	245

```

Db      62 TTTGTATCCAAAGGGGCGATCCAGTTCAAAAGGCTTCACAGGTTGCCCCCTGCGCACCTGCG 121
Qy      246 CTACCCGACCTGTGACCTCTTACCCAGCCCTGAGCGACGACCTGCTCCCATCCAG 305
Db      122 CTACCCGACCTGTGACCTCTTACCCAGCCCTGAGCGACGACCTGCTCCCATCCAG 181
Qy      306 ATCCCGCAGCCCTTGGGGGCTCCACCGGACGCACTTCCCGGGGGAATTCTGATG 365
Db      182 ATCCCGCAGCCCTTGGGGGCTCCACCGGACGCACTTCCCGGGGGAATTCTGATG 241
Qy      366 TCGCAACAGTGTGCGGAGCTACGAGAACGAGAAACGAGCTGTAGAGATGAGAGGA 425
Db      242 TCGCAACAGTGTGCGGAGCTACGAGAACGAGAAACGAGCTGTAGAGATGAGAGGA 301
Qy      426 TGAAGAGAGCTATACAAACCCAGGCTACCTGGTGTGTTCTGACAGCACCCCGGCGAC 485
Db      302 TGAAGAGAGCTATACAAACCCAGGCTACCTGGTGTGTTCTGACAGCACCCCGGCGAC 361
Qy      486 TAGCAGCTGTGCCCCCATCAGCTCCTGCACTCAGCACCCCTGCGACAGTGCCTT 545
Db      362 TAGCAGCTGTGCCCCCATCAGCTCCTGCACTCAGCACCCCTGCGACAGTGCCTT 421
Qy      546 CTCCATGAGTGCATTAATGATTAAGTGAAGCTTCCGAGAGCGGGGAGAGCGAGAAC 605
Db      422 CTCCATGAGTGCATTAATGATTAAGTGAAGCTTCCGAGAGCGGGGAGAGCGAGAAC 481
Qy      606 GTCTCTGATGAGCGAGCGGGAGTATGTAATGTGTCAGAGAACTGCATCCTGAGGGG 665
Db      482 GTCTCTGATGAGCGAGCGGGAGTATGTAATGTGTCAGAGAACTGCATCCTGAGGGG 541
Qy      666 TAAAGCTGAGGCTCCCGCCCTGAGTTCACAGAGAGAGAGAGTGAAGAGAGGGGCG 725
Db      542 TAAAGCTGAGGCTCCCGCCCTGAGTTCACAGAGAGAGAGAGTGAAGAGAGGGGCG 601
Qy      726 TCAGATTACGAGAAATCTGACGAGAGCTGAAGTGA 759
Db      602 TCAGATTACGAGAAATCTGACGAGAGCTGAAGTGA 635

```

```

RESULT 10
AL580592/c 937 bp mRNA linear EST 07-APR-2004
LOCUS
DEFINITION
Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
Homo sapiens cDNA clone CS0DJ015YE16 3-PRIME, mRNA sequence.
AL580592
AL580592.1 GI:12946762
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 937)
L1,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by life technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
8019.F

```

For more information about this cluster, see
<http://www.genoscope.cns.fr/cdnas=CS0DJ015BC08NP1&c=8019.f>.

FEATURES

source

```

+1..937
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ015YE16"

```

```

/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

Query Match 59.0%; Score 625.6; DB 1; Length 937;
 Best Local Similarity 98.3%; Pred. No. 2.9e-135;
 Matches 625; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

```

Qy      394 GAGGAACACAGCTGTGAGATGACAGATGAGAGAGAGAGAGATATCAACACAGAGCTAC 453
Db      937 GAGGAACACAGCTGTGAGATGAGAGATGAGAGAGAGAGAGATATCAACACAGAGCTAC 878
Qy      454 CTGGTGTGCTTCTCTGACAGACCCCGGCGACATGACATGCTGCCCCATGAGCTCTGCA 513
Db      877 CTGGTGTGCTTCTCTGACAGACCCCGGCGACATGACATGCTGCCCCATGAGCTCTGCA 818
Qy      514 CTGACACCCCTTGACATCCGAGACAGTGCCTTCTCCATGAGATGATTAATTAAGT 573
Db      817 CTGACACCCCTTGACATCCGAGACAGTGCCTTCTCCATGAGATGATTAATTAAGT 758
Qy      574 AAGCTTCCGAGAGCGGGGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 633
Db      757 AAGCTTCCGAGAGCGGGGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 698
Qy      634 AATGTGTCCAGAACTGCATCTGTGAGCGGCTGAAGCTGAGAGCTGAGCGGCTGAGTTCC 693
Db      697 AATGTGTCCAGAACTGCATCTGTGAGCGGCTGAAGCTGAGAGCTGAGCGGCTGAGTTCC 638
Qy      694 CAGAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 753
Db      637 CAGAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 578
Qy      754 AACTGAGGCGCTGTGAGAGCGGAGTCTGTCTGGAACAGAGCTTGTGCGGAGAGAG 813
Db      577 AACTGAGGCGCTGTGAGAGCGGAGTCTGTCTGGAACAGAGCTTGTGCGGAGAGAG 518
Qy      814 CTGGGAGCTGGAAGTGGCTCTGGGGTCTCTACATGGCGTCTGCTGCTGCTGAGCTG 873
Db      517 CTGGGAGCTGGAAGTGGCTCTGGGGTCTCTACATGGCGTCTGCTGCTGCTGAGCTG 458
Qy      874 ACAACAGCTGGAAGAAATCCCGGCTAATTAATTAATTAATTAATTAATTAATTAAT 933
Db      457 ACAACAGCTGGAAGAAATCCCGGCTAATTAATTAATTAATTAATTAATTAATTAAT 398
Qy      934 CCGAAGCTCTGACCTTGTGACGACGCTGAGATGACCTGCGGCGGCGGCGGCGGCTAC 993
Db      397 CCGAAGCTCTGACCTTGTGACGACGCTGAGATGACCTGCGGCGGCGGCGGCGGCTAC 338
Qy      994 TCTGTGTAATGATAAAGGCTGCGTGTGCTGTG 1029
Db      337 TCTGTGTAATGATAAAGGCTGCGTGTGCTGTG 302

```

```

RESULT 11
AL580240/c 922 bp mRNA linear EST 07-APR-2004
LOCUS
DEFINITION
Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
Homo sapiens cDNA clone CS0DJ005YC09 3-PRIME, mRNA sequence.
AL580240
AL580240.3 GI:46258991
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 922)
L1,W.B., Gruber,C., Jessee,J. and Polayes,D.

```


1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo. "

Query Match 57.8%; Score 612.2; DB 2; Length 651;
Best Local Similarity 98.3%; Pred. No. 3.7e-132;
Matches 625; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 385 TAGAGAAACGAGAAACGAGCTGTGAGATGAGATGAGAGACATATACACAC 444
DB 651 TAGAGAAACGAGAAACGAGCTGTGAGATGAGATGAGAGACATATACACAC 592
QY 445 CCA-GGCTACCTGTGTGTCTTCTGACAGCAACCCCGGCACTAGACCTGTGCCCCATC 503
DB 591 CCAAGGCTACCTGTGTGTGTCTTCTGACAGCAACCCCGGCACTAGACCTGTGCCCCATC 532
QY 504 AGCTCTGCACTGACACACCTTGGCATTCGAGACATGTCCTTCTCCATGAGTCCATTGA 563
DB 531 AGCTCTGCACTGACACACCTTGGCATTCGAGACATGTCCTTCTCCATGAGTCCATTGA 472
QY 564 TGATTACGTGAACGTTCCGAGAGCGGAGAGAGCGACAGACGTTCTGTGATGGACGCG 623
DB 471 TGATTACGTGAACGTTCCGAGAGCGGAGAGAGCGACAGACGTTCTGTGATGGACGCG 412
QY 624 GAGATATGTGAATGTGTCCAGAACTGCACTCTGAGCGGCTAAGACTGAGCCTGCGGC 683
DB 411 GAGATATGTGAATGTGTCCAGAACTGCACTCTGAGCGGCTAAGACTGAGCCTGCGGC 352
QY 684 CTGTAGTTCCAGAGAGGAG 743
DB 351 CTGTAGTTCCAGAGAGGAG 292
QY 744 GCAGAGAGTGAACCTGAGAGGAGGCTGTGAGAGCGGAGCTGTCTGAGAACAGAGCTTGCCCTG 803
DB 291 GCAGAGAGTGAACCTGAGAGGAGGCTGTGAGAGCGGAGCTGTCTGAGAACAGAGCTTGCCCTG 232
QY 804 GAGCGCTGAGCTGTGAG 863
DB 231 GAGCGCTGAGCTGTGAG 172
QY 864 CTGAGAGCTGACACAGGCTGAGAAATCTCCCGGTAATTTATCACTTTGGGGTTGCG 923
DB 171 CTGAGAGCTGACACAGGCTGAGAAATCTCCCGGTAATTTATCACTTTGGGGTTGCG 112
QY 924 CTGTGTGTCGCGGAGAGCTGTGACCTTCTGAGAGAGCTGAGATGACCTGCGCTGCGC 983
DB 111 CTGTGTGTCGCGGAGAGCTGTGACCTTCTGAGAGAGCTGAGATGACCTGCGCTGCGC 52
QY 984 CCAAGCCTACTCTGTGTATGAAATTAAGGCC 1015
DB 51 CCAAGCCTACTCTGTGTATGAAATTAAGGCC 20

RESULT 13
AL558866 878 bp mRNA linear EST 02-APR-2004
LOCUS AL558866 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
DEFINITION AL558866 Homo sapiens cDNA clone CS0DJ015YE16 5-PRIME, mRNA sequence.
ACCESSION AL558866
VERSION AL558866.3 GI:46184253
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 878)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:31282999.
Contact: Genoscope - Centre National de Séquençage
BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
8019.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?b=CS0DJ015BC08P1&c=8019.f.
Location/Qualifiers

FEATURES

1..878
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ015YE16"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_11b="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 57.7%; Score 612; DB 1; Length 878;
Best Local Similarity 98.3%; Pred. No. 4.3e-132;
Matches 631; Conservative 7; Mismatches 2; Indels 2; Gaps 2;

QY 1 GACTTGCCTTGTGAGGGGCTTAGGGGTGACAGCCAGCTGTCTCCGAGCTCCCTGCAATG 60
DB 239 GACTTGCCTTGTGAGGGGCTTAGGGGTGACAGCCAGCTGTCTCCGAGCTCCCTGCAATG 298
QY 61 GAGAGGCAATCCGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
DB 299 GAGAGGCAATCCGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 358
QY 121 TTGATGCACTGTGTGTGCACTGCGACAGACTGCGAGCTCTTACAGACAGACATCTCTCA 180
DB 359 TTGATGCACTGTGTGTGCACTGCGACAGACTGCGAGCTCTTACAGACAGACATCTCTCA 418
QY 181 GATGTTTGTATCCAGAGGGGATTCAGATTCAACAGGCTCTACACAGTTGGCCCTTGCCCA 240
DB 419 GATGTTTGTATCCAGAGGGGATTCAGATTCAACAGGCTCTACACAGTTGGCCCTTGCCCA 478
QY 241 CTGCTACCACTGTGACCTCTCACTCAACCCCTGAGCCAGGCGAGACTGTCTCCCATC 300
DB 479 CTGCTACCACTGTGACCTCTCACTCAACCCCTGAGCCAGGCGAGACTGTCTCCCATC 538
QY 301 CCAAGATCCCGCAGCCCTTGTGGGGCTCCACCGGAGCGCATTTTCCGGCGGGAATTCT 360
DB 539 CCAAGATCCCGCAGCCCTTGTGGGGCTCCACCGGAGCGCATTTTCCGGCGGGAATTCT 598
QY 361 GATGTCGCAACAGTGTGCGAGCTTACGAGAACGAGAACGAGCTGTGAGATGCAAT 420
DB 599 SATGCT-SCAACAATGTGCGAGCTTACGAGAACGAGAACGAGCTGTGAGATGCAAT 657
QY 421 GAGATGAGGACGACTTACCAACCCAGGCTACCTGTGTGTGCTTCTCTGACAGACCCCG 480
DB 658 GAGATGAGGACGACTTACCAACCCAGGCTACCTGTGTGTGCTTCTCTGACAGACCCCG 717
QY 481 GCCATGACACTGTGCGCCCATCAGTCTGTGACTGACACCCCTGTGCATCCGAGACGT 540
DB 718 GSCATGACACTGTGCGCCCATCAGTCTGTGACTGACACCCCTGTGCATCCGAGACGT 777
QY 541 GCCTTCTCCATGAGTCCATTGATGATTAAGTGAACCTTCCGAGAGCGGGAGAGCGCA 600
DB 778 GCCTTCTCCATGAGTCCATTGATGATTAAGTGAACCTTCCGAGAGCGGGAGAGCGCA 836
QY 601 GAAGCTCTCTGATGTGACCGGAGATGATGTAATGTGCC 642
DB 837 GAAGCTCTCTGATGTGACCGGAGATGATGTAATGTGCC 878

RESULT 14
LOCUS BM715456 610 bp mRNA linear EST 28-FEB-2002
DEFINITION UI-E-EJ0-ah1-g-12-0-UI.r2 UI-E-EJ0 Homo sapiens cDNA clone
BM715456
ACCESSION BM715456
VERSION BM715456.1 GI:19028714
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 610)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.
Location/Qualifiers
1. 610
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-EJ0-ah1-g-12-0-UI"
/tissue_type="fetal eyes, lens, eye anterior segment,
optic nerve, retina, Retina Foveal and Macular, RPE and
Choroid"
/dev_stage="fetal and adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_1lb="UI-E-EJ0"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-EJ0 is a subcloned cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for this library are: fetal eyes,
AGATTCACAGA; lens, CGATTACCA; eye anterior segment,
AATGCCCAT; optic nerve, CCATTAGTG; retina, CCGCG; Retina
Foveal and Macular, GTCC; RPE and Choroid, ACCCA. This
library was created for the program, Gene Discovery in the
Visual System, supported by National Eye Institute (NEI)."

ORIGIN
Query Match 57.2%; Score 606.4; DB 4; Length 610;
Best Local Similarity 99.5%; Pred. No. 8.3e-131;
Matches 607; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 416 CAGATGAGATGAGACGATCAACCCAGGCTACCTGGTGCTTCTGACACAGA 475
DB 1 CGGATGAGATGAGACGATCAACCCAGGCTACCTGGTGCTTCTGACACAGA 60

QY 476 CCGCGGACCTAGACATGCTGCTGCCCCCATGACCTCTGACCTGACACACCCCTGGCATCCGAG 535
DB 61 CCGCGGACCTAGACATGCTGCTGCCCCCATGACCTCTGACCTGACACACCCCTGGCATCCGAG 120
QY 536 ACAGTGCCTTCTCCATGAGAGTCCATTTGATGATTAAGTGAACGTTCCGAGACGGGGGAGA 595
DB 121 ACAGTGCCTTCTCCATGAGAGTCCATTTGATGATTAAGTGAACGTTCCGAGACGGGGGAGA 180
QY 596 GCGCAGAAAGCGTCTTGATGAGCAGCGGAGATATGTAATGTCTCCAGCAATCTGCATC 655
DB 181 GCGCAGAAAGCGTCTTGATGAGCAGCGGAGATATGTAATGTCTCCAGCAATCTGCATC 240
QY 656 CTGAGCGGCTTGAAGCTGAGCTGCGGCTGAGTTCCAGAGGACAGAGAAAGTGAAG 715
DB 241 CTGAGCGGCTTGAAGCTGAGCTGCGGCTGAGTTCCAGAGGACAGAGAAAGTGAAG 300
QY 716 AAGAGGGGCTTCCATTAACGAAATCTGACAGAGACTGAAGAGGCTGAGAGGCG 775
DB 301 AAGAGGGGCTTCCATTAACGAAATCTGACAGAGACTGAAGAGGCTGAGAGGCG 360
QY 776 AGTCTGTCTTGAACCAAGCTTGGCTGGAGAGGCTGAGCTGGAGCACTGGAAGTGGCT 835
DB 361 AGTCTGTCTTGAACCAAGCTTGGCTGGAGAGGCTGAGCTGGAGCACTGGAAGTGGCT 420
QY 836 GGGGTCTTCAATGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 895
DB 421 GGGGTCTTCAATGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 896 CGTAACTATTATCACTTTGGGGTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 955
DB 481 CGTAACTATTATCACTTTGGGGTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 956 CGCAGCCTGAGAAATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1015
DB 541 CGCAGCCTGAGAAATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 1016 TGGCTGTGTC 1025
DB 601 TGGCTGTGTC 610

RESULT 15
LOCUS BX437371 811 bp mRNA linear EST 04-MAY-2004
DEFINITION BX437371 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP006YF09
5-PRIME, mRNA sequence.
ACCESSION BX437371
VERSION BX437371.2 GI:47003573
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 811)
Li, W.B., Gruber, C., Jeesee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
COMMENT On May 15, 2003 this sequence version replaced gi:30777562.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 8019.f.
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?S=CS0CAP006C05Qp1&c=8019.f.
FEATURES
source
1. 811
/organism="Homo sapiens"

```
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/clone="CS0CAP006YF09"  
/issue_type="THYMUS"  
/clone_lib="Homo sapiens THYMUS"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo(dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."
```

ORIGIN

```
Query Match      54.5%; Score 578; DB 5; Length 811;  
Best Local Similarity 100.0%; Pred. No. 3,7e-124;  
Matches 578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY      1 GACTCTGCGCTTTGAGGGGCTTAGGGGTGCAGCCAGCTTCTCCGAGCTCCCTGCAAGT 60  
        |||  
Db      234 GACTCTGCGCTTTGAGGGGCTTAGGGGTGCAGCCAGCTTCTCCGAGCTCCCTGCAAGT 293  
  
QY      61 GAGAGGCGCATCTGTGCTCCCTGCGTGGGGCTCTGTGCTGCTGCCATCTGGGCATG 120  
        |||  
Db      294 GAGAGGCGCATCTGTGCTCCCTGCGTGGGGCTCTGTGCTGCTGCCATCTGGGCATG 353  
  
QY      121 TTGATGGCATGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
        |||  
Db      354 TTGATGGCATGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 413  
  
QY      181 GATAGTTGTATCCAAAGGGGATCCAGTTCAAAGGCGCTCAAGGTTGCGCCCTGGGCA 240  
        |||  
Db      414 GATAGTTGTATCCAAAGGGGATCCAGTTCAAAGGCGCTCAAGGTTGCGCCCTGGGCA 473  
  
QY      241 CCTGCTTACCCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
        |||  
Db      474 CCTGCTTACCCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 533  
  
QY      301 CCAAGATCCCGCAAGCCCTTTGGGGGCTCCCAACGAGCGCATTTCCGGGCGGATTCT 360  
        |||  
Db      534 CCAAGATCCCGCAAGCCCTTTGGGGGCTCCCAACGAGCGCATTTCCGGGCGGATTCT 593  
  
QY      361 GATGTGCAACAGTGTGGCGAGCTACGAAACGAGAAACGAGCTGTGAGATGCAAT 420  
        |||  
Db      594 GATGTGCAACAGTGTGGCGAGCTACGAAACGAGAAACGAGCTGTGAGATGCAAT 653  
  
QY      421 GAGATGAGAGCACTATCACAAACGAGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
        |||  
Db      654 GAGATGAGAGCACTATCACAAACGAGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 713  
  
QY      481 GCCACTAGCACTGTGCGCCATCAGCTCTGCACTCAGACCCCTGGCATCCGAGCAGT 540  
        |||  
Db      714 GCCACTAGCACTGTGCGCCATCAGCTCTGCACTCAGACCCCTGGCATCCGAGCAGT 773  
  
QY      541 GCCTTTCCATGAGAGTGCATTGATGATTAAGTGAACGT 578  
        |||  
Db      774 GCCTTTCCATGAGAGTGCATTGATGATTAAGTGAACGT 811
```

Search completed: November 22, 2004, 12:04:48
Job time : 6720 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 14:06:44 ; Search time 155 Seconds
(without alignments)
469.820 Million cell updates/sec

Title: US-09-597-920B-4_COPY_31_233
Perfect score: 1071

Sequence: 1 RLPSYDSTSSDSLPRGIQ.....EAEVEERGAPDYENLQELN 203

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_23Sep04: *
1: geneseqp1980s: *
2: geneseqp1990s: *
3: geneseqp2000s: *
4: geneseqp2001s: *
5: geneseqp2002s: *
6: geneseqp2003as: *
7: geneseqp2003bs: *
8: geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1071	100.0	233	2	AAV27120 Human LAT
2	683	63.8	242	2	AAV27121 Murine LAT
3	132	12.3	710	5	ABB91536 Herbicida
4	128	12.0	374	6	ABP72770 Thiorodox
5	122	11.4	704	7	AD660099 Rat Prote
6	122	11.4	704	7	ADDA7584 Rat Prote
7	122	11.4	704	7	AD660734 Rat Prote
8	122	11.4	704	7	AD660097 Rat Prote
9	122	11.4	704	7	ADDA7582 Rat Prote
10	122	11.4	704	7	ADDA7582 Rat Prote
11	121	11.3	699	7	ADDA7582 Rat Prote
12	121	11.3	718	4	AAW80140 Human pro
13	121	11.3	742	8	ADP04241 Human col
14	116	10.8	2058	4	AAW80140 Human col
15	116	10.8	2127	7	ADP09542 Human pol
16	114.5	10.7	838	4	ADP09542 Human pol
17	114.5	10.7	1000	6	ABW70548 Drosophill
18	113.5	10.6	735	6	ABW25647 Aspergill
19	113.5	10.6	1186	4	ABU41468 Protein e
20	112.5	10.5	1177	6	ABW60992 Drosophill
21	112.5	10.5	1186	6	ABP71253 Chimeric
22	112	10.5	519	4	ABP71253 Zebrafish
23	112	10.5	519	4	ABW90786 Human she
24	112	10.5	917	6	ABW99466 Amino aci
25	112	10.5	923	6	ADA13333 Human int
					ABW54617 Human NOV

26	112	10.5	1958	2	AAW60620 Protein f
27	111.5	10.4	216	6	ABP72769 Human nuc
28	111	10.4	1048	2	AAW27277 Human cyt
29	111	10.4	1807	4	AAW85697 Recombina
30	111	10.4	2028	4	AAW85698 Recombina
31	110.5	10.3	348	4	ABW69117 Drosophill
32	110	10.3	505	5	AAU77474 EVPR prot
33	110	10.3	571	7	ABW81787 Pseudomon
34	109	10.2	966	4	AAW38679 Human pol
35	109	10.2	1013	4	AAW38678 Human pol
36	109	10.2	1025	4	AAW38680 Human pol
37	108.5	10.1	774	7	ADC31626 Human nov
38	108.5	10.1	1312	7	ABP69782 Human nov
39	108.5	10.1	1343	7	AD169279 Human hea
40	108.5	10.1	1714	6	ABJ25518 Aspergill
41	108.5	10.1	1750	6	ABJ26118 Aspergill
42	108	10.1	732	4	ABW59675 Drosophill
43	108	10.1	2296	4	ABW59050 Drosophill
44	107.5	10.0	634	7	ADM04655 Human pro
45	107.5	10.0	1822	5	ABP43899 Nuclear p

ALIGNMENTS

RESULT 1	AAV27120	standard; protein; 233 AA.
ID	AAV27120	standard; protein; 233 AA.
AC	AAV27120;	
DT	14-SEP-1999	(first entry)
XX		
DE	Human LAT (linker for activation of T cells) protein.	
KW	LAT; tyrosine kinase; linker for activation of T cell; TCR; human; T-cell receptor; TCR signalling pathway; neoplasia; inflammation; hypersensitivity; allergy; microbial infection; genetic disease; autoimmune disease; graft rejection; modulator.	
KW	Homo sapiens.	
OS	W09932627-A2.	
PN	01-JUL-1999.	
PD	23-DEC-1997;	98WO-US027400.
XX		
PR	23-DEC-1997;	97US-0068690P.
XX		
XX	(USSH) US DEPT HEALTH & HUMAN SERVICES.	
XX	Samelson LE, Zhang W;	
DR	WPI. 1999-418926/35.	
DR	N-PSDB; AAX89072.	
XX		
PT	Linker for activation of T cell protein used to, e.g. screen for modulators of T cell signaling.	
PT	Claim 1; Fig 7C; 125bp; English.	
PS	The invention relates to a protein tyrosine kinase substrate LAT (linker for activation of T cells) protein. Modulation of interaction between LAT and the T-cell receptor (TCR) affects the TCR signalling pathway. LAT is a substrate for tyrosine kinases and becomes phosphorylated after TCR engagement, resulting in recruitment of other signalling molecules. LAT is used to identify and test (ant)agonists of tyrosine kinase signalling pathways, i.e. modulation of interaction between tyrosine kinase substrates and intracellular ligands or between these ligands and other members of the pathway, including identification of downstream signalling proteins, particularly in immune system cells. These modulators are potentially useful as drugs and diagnostic agents, particularly for	

CC diseases that involve undesirable cell proliferation, differentiation,
 CC growth or T cell anergy, e.g. neoplasia, inflammation, hypersensitivity/
 CC allergy, microbial infection, metabolic, genetic or autoimmune diseases,
 CC graft rejection. LAT is also used to generate specific antibodies, used
 CC for detection of LAT. Nucleic acid that encodes LAT, or its fragments,
 CC are used to identify homologous sequences in other species; to detect the
 CC LAT gene and as sources of antisense therapeutics. Modulators of LAT are
 CC potentially more specific and less toxic than known immunosuppressants
 CC such as cyclosporin. The present sequence represents the amino acid
 CC sequence of human LAT

XX
 SQ Sequence 233 AA;

Query Match 100.0%; Score 1071; DB 2; Length 233;
 Best Local Similarity 100.0%; Pred. No. 6,2e-84;
 Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLPGSYDSTSSDLYPRGIGQKRPHTVAPWPPAYPVTSYPLSQPDLPIRSPQPLG 60
 DB 31 RLPGSYDSTSSDLYPRGIGQKRPHTVAPWPPAYPVTSYPLSQPDLPIRSPQPLG 90

QY 61 SHRTPSRRSDSGANSVASYENEPACDADDEDYHNPGLVLPDSTPATSTAAPS 120
 DB 91 SHRTPSRRSDSGANSVASYENEPACDADDEDYHNPGLVLPDSTPATSTAAPS 150

QY 121 PALSTPGIRDSAFMESIDYVNVPESGESAASLDGSRVYVNSQELHPGAATPEPAL 180
 DB 151 PALSTPGIRDSAFMESIDYVNVPESGESAASLDGSRVYVNSQELHPGAATPEPAL 210

QY 181 SSQAEEVEEKGADYENLQELN 203
 DB 211 SSQAEEVEEKGADYENLQELN 233

RESULT 2
 AAY27121
 ID AAY27121 standard; protein; 242 AA.

XX
 AC AAY27121;
 XX
 DT 14-SEP-1999 (first entry)

XX
 DE Murine LAT (linker for activation of T cells) protein.

XX
 KM LAT; tyrosine kinase; linker for activation of T cell; TCR; human;
 KM T-cell receptor; TCR signalling pathway; neoplasia; inflammation;
 KM hypersensitivity; allergy; microbial infection; genetic disease;
 KM autoimmune disease; graft rejection; modulator; mouse.

XX
 OS Mus musculus.

XX
 PN WO9932627-A2.

XX
 PD 01-JUL-1999.

XX
 PF 23-DEC-1998; 98WO-US027400.

XX
 PR 23-DEC-1997; 97US-0068690P.

XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX
 PI SameLson LE, Zhang W;
 XX
 DR WPI; 1999-418926/35.
 XX
 DR N-PSDB; AAX89074.

XX
 PT Linker for activation of T cell protein used to, e.g. screen for
 XX
 PS modulators of T cell signaling.
 XX
 PS Example 2; Fig 7D; 125pp; English.

XX
 CC The invention relates to a protein tyrosine kinase substrate LAT (linker
 CC for activation of T cells) protein. Modulation of interaction between LAT

CC and the T-cell receptor (TCR) affects the TCR signalling pathway. LAT is
 CC a substrate for tyrosine kinases and becomes phosphorylated after TCR
 CC engagement, resulting in recruitment of other signalling molecules. LAT
 CC is used to identify and test (ant)agonists of tyrosine kinase signalling
 CC pathways, i.e. modulation of interaction between tyrosine kinase
 CC substrates and intracellular ligands or between these ligands and other
 CC members of the pathway, including identification of downstream signalling
 CC proteins, particularly in immune system cells. These modulators are
 CC potentially useful as drugs and diagnostic agents, particularly for
 CC diseases that involve undesirable cell proliferation, differentiation,
 CC growth or T cell anergy, e.g. neoplasia, inflammation, hypersensitivity/
 CC allergy, microbial infection, metabolic, genetic or autoimmune diseases,
 CC graft rejection. LAT is also used to generate specific antibodies, used
 CC for detection of LAT. Nucleic acid that encodes LAT, or its fragments,
 CC are used to identify homologous sequences in other species; to detect the
 CC LAT gene and as sources of antisense therapeutics. Modulators of LAT are
 CC potentially more specific and less toxic than known immunosuppressants
 CC such as cyclosporin. The present sequence represents the amino acid
 CC sequence of murine LAT

XX
 SQ Sequence 242 AA;

Query Match 63.8%; Score 683; DB 2; Length 242;
 Best Local Similarity 68.9%; Pred. No. 1.5e-50;
 Matches 146; Conservative 16; Mismatches 38; Indels 12; Gaps 6;

QY 2 LPGSYDSTSSDLYPRGIGQKRPHTVAPWPPA--YPPVTSYPLSQPDLPIRSPQPLG 59
 DB 33 LPVSYDSTSSDLYPRGIGQKRPHTVAPWPPA--YPPVTSYPLSQPDLPIRSPQPLG 92

QY 60 GSHRTPSRRSDSGANSVASYENEPACDADDEDYHNPGLVLPDSTPATSTA 117
 DB 93 GSHRTPSRRSDSGANSVASYENEPACDADDEDYHNPGLVLPDSTPATSTA 151

QY 118 PSAPALSTPGIRDSAFMESIDYVNVPESGESAASLDGSRVYVNSQELHPGAATPE 177
 DB 152 PSAPALSTPGIRDSAFMESIDYVNVPESGESAASLDGSRVYVNSQELHPGAATPE 210

QY 178 AALSSQEA--EVEEKGADYENLQELN 203
 DB 211 AALSSQEA--EVEEKGADYENLQELN 242

RESULT 3
 ABB91536
 ID ABB91536 standard; protein; 710 AA.

XX
 AC ABB91536;
 XX
 DT 31-MAY-2002 (first entry)

XX
 DE Herbicidally active polypeptide SEQ ID NO 747.

XX
 KM Herbicidal; plant; agriculture; herbicide.

XX
 KM Arabidopsis thaliana.

XX
 OS Arabidopsis thaliana.

XX
 PN WO200210210-A2.

XX
 PD 07-FEB-2002.

XX
 PF 28-AUG-2001; 2001WO-EP009892.

XX
 PR 28-AUG-2001; 2001WO-EP009892.

XX
 PA (FARB) BAYER AG.

XX
 PI Tietjen K, Weidler M;
 XX
 DR WPI; 2002-269010/31.

XX
 PT Identifying plant target proteins for herbicidally active compounds,
 XX
 CC comprising aligning and comparing nucleic acid or amino acid sequences

PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms.
XX
PS Claim 5; SEQ ID NO 747; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins (AB990790-AB94016)
CC for herbicidally active compounds, comprising aligning and comparing
CC nucleic acid or amino acid sequences from plant with nucleic acid or
CC amino acid sequences from non-plant organisms using suitable search
CC parameters, where plant sequences having an E-value greater by a factor
CC of 3 than the E-value of most similar non-plant sequences are selected.
CC The polypeptides or nucleic acids encoding them are useful for
CC identifying modulators. The identified modulators are useful as
CC herbicides
XX
SQ Sequence 710 AA;

Query Match 12.3%; Score 132; DB 5; Length 710;
Best Local Similarity 31.6%; Pred. No. 0.015;
Matches 49; Conservative 6; Mismatches 52; Indels 48; Gaps 6;

QY 8 STSDSLYPRGIQ-----FKRPTVAPWPVPTSYPLSQPDLPIPRSPQLGGS 61
DB 40 SPPADSSPPALPSPAPVFPSPPTVSSPP--PPLDSSP--PPDLTRFPSSPPPPAP 96
QY 62 -----HRTPSRRSDGANSVASYENEPACEDADEDED----- 95
DB 97 PPPIVFPPIIDSPPESTNSPPPEVEFPPEPPADEDESPAPPPPEQLPPPASSPOG 156
QY 96 ----DYHNPGLVLPDSTPATSTAPASAPALSTP 126
DB 157 PKKPKGHRG-----PATSPAPSPATSP 182

RESULT 4
ABP72770
ID ABP72770 standard; protein; 374 AA.
XX
AC ABP72770;
XX
DT 23-OCT-2003 (revised)
DT 11-JUN-2003 (first entry)
XX
DE Thioresoxin-hNop140 C-terminal region fusion protein.
XX
KW Human; nucleolar phosphoprotein; hNop140; Trx-hNop140C; doxorubicin;
KW cytosolic; thioresoxin.
XX
OS Homo sapiens.
OS Escherichia coli.
OS Chimeric.
XX
PN MO2003011904-A1.
XX
PD 13-FEB-2003.
XX
PF 19-OCT-2001; 2001WO-KR001771.
XX
PR 30-JUL-2001; 2001KR-00045860.
XX
PA (KOAD) KOREA ADV INST SCI & TECHNOLOGY.
XX
PI Yu Y, Jin Y, Yu J;
XX
DR WPI; 2003-278388/27.
XX
N-PSDB; AB282080.
XX
PT New micro-terminal region protein of human nucleolar phosphoprotein
PT hNop140, useful for reducing the side effects and improving activity of
PT an anticancer agent.
XX
PS Claim 6; Page 31-33; 36pp; English.
XX

CC The present sequence is the protein sequence of a fusion protein, termed
CC Trx-hNop140C, comprising thioresoxin linked to the N-terminus of the C-
CC terminal region (see also ABP72769) of human nucleolar phosphoprotein
CC hNop140. The C-terminal region of hNop140 was identified by biopanning
CC as an intracellular target of the anticancer agent, doxorubicin. cDNA
CC encoding the C-terminal region was amplified by PCR and inserted into
CC pET28a so as to produce the Trx-hNop140C fusion protein, which was
CC obtained in large quantities in Escherichia coli transformants. The
CC fusion protein retains the original form of hNop140 C-terminal region,
CC and is phosphorylated by casein kinase. The C-terminal region of hNop140
CC binds to doxorubicin within cells, thus playing an important role in
CC mediating the cellular toxicity or anticancer activity of doxorubicin.
CC The hNop140 C-terminal region is expected to be useful for reducing the
CC side effects of doxorubicin and for improving the anticancer activity of
CC doxorubicin. (Updated on 23-OCT-2003 to standardise OS field)
XX
SQ Sequence 374 AA;

Query Match 12.0%; Score 128; DB 6; Length 374;
Best Local Similarity 26.5%; Pred. No. 0.015;
Matches 53; Conservative 24; Mismatches 87; Indels 36; Gaps 7;

QY 4 GSYDSTSDSLYPRGIQPKRPTVAPWPVPTSYPLSQPDLPIPRSPQLGSHR 63
DB 132 GSGDDDDKVLVPRGSMGNPNS-----SNKPAVTTKPAVKPAAP----KQVGGGCK 182
QY 64 TPSSRRSDGANSVASYENEPACEDADEDDYHNPGLVLPDSTPATSTAPSPAL 123
DB 183 ILTRRADSSSEESSSEEEKTK-----WVATTKPATATAALSLPAK 227
QY 124 STP-CIRPSAFMSIEDYVNVPEGSGAASLDSREYVNV--OELRPGAKTPAPALS 181
DB 228 QAPQSSRDS--SSDS-----DSSSESEEEKTKSAVKKKPKQVAGAAPSPASAK 277
QY 182 SQAEVEEGAPDYENLOE 201
DB 278 KKAESSNSSSSDDSEEBE 297

RESULT 5
ADE60099
ID ADE60099 standard; protein; 704 AA.
XX
AC ADE60099;
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein P41777, SEQ ID NO 5998.
XX
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
XX
PN MO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
XX
PR 01-NOV-2001; 2001US-0346382P.
XX
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GHEO) GEN HOSPITAL CORP.
XX
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
XX
DR GENBANK; F41777.
XX
PT New composition comprising two or more isolated polypeptides, useful for

PD 27-FEB-2003

AD660734

```

ID AD60734standard; protein; 704 AA.
AC ADE60734;
XX
XX 29-JAN-2004 (first entry)
XX
XX Rat Protein P41777, SEQ ID NO 6646.
XX
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
XX Rattus norvegicus.
XX
XX WO2003016475-A2.
XX
XX 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
XX
XX 01-NOV-2001; 2001US-0346382P.
XX
XX 26-NOV-2001; 2001US-0333347P.
XX
XX (GENO ) GEN HOSPITAL CORP.
XX (FARB ) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX
XX MPI; 2003-268312/26.
XX
XX GENBANK; P41777.
XX
XX New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the
XX specification, a method for identifying a compound useful in treating
XX pain and a pharmaceutical composition comprising the one or more
XX polypeptides or their antibodies. The polynucleotide or the compound that
XX modulates its activity is useful for preparing a medicament for treating
XX pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX therapy). The sequence presented is a rat protein (shown in Table 2 of
XX the specification) which is differentially expressed during pain. Note:
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic form directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 704 AA;
XX
XX Query Match 11.4%; Score 122; DB 7; Length 704;
XX Best Local Similarity 23.2%; Pred. No. 0.11;
XX Matches 54; Conservative 25; Mismatches 88; Indels 66; Gaps 6
XX
XX 5 SYDSTSSSSSLVGRGIQFKRPHVAVPMPVAVPYPTSTSPVLSQDPLALIP-----RSPQL 58
XX |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 372 SSDSSDSSSDEDA-----PAKKPVSATSSPLSKPVAVTRKPPAKAVAVTRKOP 418

```

Oy	55	CGSHRPPSSRRDDGANSVASYENPEPCCEADDEBDYHNPGVLVLPDSTPATSTRAAP	116	
Db	419	AGSGGQRPQSKKADSSSESESSSSSEBEATKKS	-----YTPKRVYTAQAP	464
Oy	119	SAPALSTP---GIRDSAFMGESIDYVWVP	-----ESGES	156
Db	465	SLRPAKQAPRAGGSSSDSSSSSEBEKKTTPRPAPKKAAGAAPVKTPTPVKAAAESSS	524	
Oy	151	AEASLDGSRN-----YVNVSQELHFGAAKTEBPALSSCEAEVEBEGAPDYN	198	
Db	525	SSSESDSESEBEKKPKSKATPPRQAKANGVAPASONGRAGRESEEBEDTEON	577	
RESULT 8				
AD660097	ID	AD660097 standard; protein; 704 AA.		
XX	AC	AD660097;		
XX	DT	29-JAN-2004 (first entry)		
XX	DE	Rat Protein P41777, SEQ ID NO 5996.		
KW	RA	Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;		
KX	CH	chronic constriction injury; CCI; spared nerve injury; SNI; Chung.		
OS	NC	Rattus norvegicus.		
XX	PN	MO2003016475-A2.		
PD	27-FEB-2003.			
XX	PF	14-AUG-2002; 2002MO-US025765.		
PR	14-AUG-2001; 2001US-0312147P.			
PR	01-NOV-2001; 2001US-0346382P.			
PR	26-NOV-2001; 2001US-033347P.			
PA	(GENO) GEN HOSPITAL CORP.			
PA	(FARB) BAYER AG.			
P1	WOOLF C, D'URSO D, BEFORT K, COSTIGAN M;			
XX	WPI; 2003-268312/26.			
DR	GENEANK; P41777.			
XX	New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.			
PS	Claim 1; Page; 1017pp; English.			
XX	The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction			

CC injury (CCI) and spared nerve injury (SNI) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 704 AA:

Query Match 11.4%; Score 122; DB 7; Length 704;
Best Local Similarity 23.2%; Pred. No. 0.11;
Matches 54; Conservative 25; Mismatches 88; Indels 66; Gaps 6;

5 SYDSTSDSLVPRGIGFQKPHHTVAPWPAPYPTVSYPLSQPDLPLP-----RSQPL 58
372 SSDDSDSDSDEA-----PAKVSATKSPKSPAVTPKPAKAVATKQP 418
59 GGSHTPTSSRRDSDGANSVASYENEPACEDADEDEDYHNPGLVLPDSTPATSTAP 118
419 AGSGQKQSRKADSSSSSESSSESEATKKS-----VTPKAVTAKAP 464
119 SAPALSTP---GIRDSAFSMESIDYVNP-----ESGES 150
465 SLPAKQAPRAGDSSSDSSSESSSESEKTPPKPAKKAAGAVPKTPVKAAAESSSS 524
151 AEASLDGSR-----YVNVSOELHPGAKTEPALSSQAEVEEAGAPYEN 198
525 SSSSESDSSEKPKSKATPKPAQKANGVPASONGKAGESEEDTEQN 577

RESULT 9

ADD47582
ID ADD47582 standard; protein; 704 AA.

AC ADD47582;

DT 29-JAN-2004 (first entry)

DE Rat Protein AAA41719, SEQ ID NO 13278.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;

KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

OS Rattus norvegicus.

PN MO2003016475-A2.

XX 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

PA (GENO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

XX Woolf C, D'Urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

DR GENBANK; AAA41719.

XX New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017p; English.

XX The invention discloses a composition comprising two or more isolated rat

CC or human polynucleotides or a polynucleotide which represents a fragment,

CC derivative of allelic variation of the nucleic acid sequence. Also

CC claimed are a vector comprising the novel polynucleotide, a host cell

CC comprising the vector, a method for identifying a nucleotide sequence

CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their analogues. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 704 AA:

Query Match 11.4%; Score 122; DB 7; Length 704;
Best Local Similarity 23.2%; Pred. No. 0.11;
Matches 54; Conservative 25; Mismatches 88; Indels 66; Gaps 6;

5 SYDSTSDSLVPRGIGFQKPHHTVAPWPAPYPTVSYPLSQPDLPLP-----RSQPL 58
372 SSDDSDSDSDEA-----PAKVSATKSPKSPAVTPKPAKKAAGAVPKTPVKAAAESSSS 524
59 GGSHTPTSSRRDSDGANSVASYENEPACEDADEDEDYHNPGLVLPDSTPATSTAP 118
419 AGSGQKQSRKADSSSSSESSSESEATKKS-----VTPKAVTAKAP 464
119 SAPALSTP---GIRDSAFSMESIDYVNP-----ESGES 150
465 SLPAKQAPRAGDSSSDSSSESSSESEKTPPKPAKKAAGAVPKTPVKAAAESSSS 524
151 AEASLDGSR-----YVNVSOELHPGAKTEPALSSQAEVEEAGAPYEN 198
525 SSSSESDSSEKPKSKATPKPAQKANGVPASONGKAGESEEDTEQN 577

RESULT 10

ADD47665
ID ADD47665 standard; protein; 704 AA.

AC ADD47665;

DT 29-JAN-2004 (first entry)

DE Rat Protein NP_074060, SEQ ID NO 13361.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;

KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

OS Rattus norvegicus.

PN MO2003016475-A2.

XX 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

PA (GENO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

Pt New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.

Fs Claim 1; Page; 1017pp; English.

Cc The invention discloses a composition comprising two or more isolated rat
Cc or human polynucleotides or a polynucleotide which represents a fragment,
Cc derivative or allelic variation of the nucleic acid sequence. Also
Cc claimed are a vector comprising the novel polynucleotide, a host cell
Cc comprising the vector, a method for identifying a nucleotide sequence
Cc which is differentially regulated in an animal subjected to pain and a
Cc kit to perform the method, an array, a method for identifying an agent
Cc that increases or decreases the expression of the polynucleotide sequences
Cc that is differentially expressed in neuronal tissue of a first animal
Cc subjected to pain, a method for identifying a compound which regulates
Cc the expression of a polynucleotide sequence which is differentially
Cc expressed in an animal subjected to pain, a method for identifying a
Cc compound that regulates the activity of one or more of the
Cc polynucleotides, a method for producing a pharmaceutical composition, a
Cc method for identifying a compound or small molecule that regulates the
Cc activity in an animal of one or more of the polypeptides given in the
Cc specification, a method for identifying a compound useful in treating
Cc pain and a pharmaceutical composition comprising the one or more
Cc polypeptides or their antibodies. The polynucleotide or the compound that
Cc modulates its activity is useful for preparing a medicament for treating
Cc pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
Cc injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
Cc therapy). The sequence presented is a rat protein (shown in Table 2 of
Cc the specification) which is differentially expressed during pain. Note:
Cc The sequence data for this patent did not form part of the printed
Cc specification, but was obtained in electronic form directly from WIPO at
Sq ftp.wipo.int/pub/published_pct_sequences.

Xx

Xx Sequence 704 AA;

Df

Query Match	11.4%;	Score 123;	DB 7;	Length 704;
Best local Similarity	23.2%;	Pred. No. 0.11;	88;	Indels 66; Gaps 6
Matches 54;	Conservative 25;	Mismatches		

Oy 5 SYDSTSSDLVPRGIGQFKRPHTVAPWPVPAYPTSYTSPQLRLRP-----RSQPFL 58
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Dd 372 SSDSDSDSSDEEA-----PAKPVSATSPSLSKPVTTRKPPAKAAVAITPKOP 418

Oy 59 GGSHTPEGRRDGDGANSVAASYENEEBPACEDADEDEDYHNPGLVLYLPDSTPATSTAAP 118
||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Dd 419 AGSGCGKPSRKADSSSEBSSEBSSEBTATKS-----VTPKKAVTKAAP 464

Oy 119 SAPALSTP---GIRSAFEMESIDYVNVP-----EGGES 150
||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Dd 465 SLPAQAQAPRAGCDDSSDSSESSEBEKKTTPPKRPARKKAAGAAPPRTPYKKAASSSSS 524

Oy 151 AEASLDGSGRE-----YVNVSQELHPAATKEPALASSQAEVEEYGAPDYEN 198
::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Dd 525 SSSSDSSBEEKKRKSKATTPKPQCKANGCVAPSNGKAGKBEBBBDTEQN 577

Rstult 11

AAM79156

ID AAM79156 standard; protein; 699 AA.

Ac AAM79156;

Dt 06-NOV-2001 (first entry)

Df Human protein SEQ ID NO 1818.

De

Km Human, cytokine; cell proliferation; cell differentiation; gene therapy;

KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KM	tissue growth factor; immunomodulatory; cancer; leukaemia;
KM	nervous system disorder; arthritis; inflammation.
XX	
OS	Homo sapiens.
PN	WO200157190-A2.
XX	
PD	09-AUG-2001.
XX	
PF	05-FEB-2001; 2001WO-US004098.
XX	
PR	03-FEB-2000; 2000US-00496914.
PR	27-APR-2000; 2000US-00560875.
PR	20-JUN-2000; 2000US-00598075.
PR	19-JUL-2000; 2000US-00620325.
PR	01-SEP-2000; 2000US-00654936.
PR	15-SEP-2000; 2000US-00663561.
PR	20-OCT-2000; 2000US-00693325.
PR	30-NOV-2000; 2000US-00728422.
XX	
PA	(HYSEQ-) HYSEQ INC.
XX	
PI	Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI	Mao Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI	Xue HJ, Yang Y, Wejhtman T, Goodrich R;
XX	
DR	WP1; 2001-476283/51.
DR	N-PSDB; AAK52289.
XX	
PT	Nucleic acids encoding polypeptides with cytokine-like activities, useful
FT	in diagnosis and gene therapy.
XX	
PS	Claim 20; Page 4189-4191; 6221IP; English.
XX	
CC	The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC	encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC	cytokine, cell proliferation or cell differentiation or which may induce
CC	production of other cytokines in other cell populations. The
CC	polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC	peptide therapy. The polypeptides have various cytokine-like activities,
CC	e.g. stem cell growth factor activity, haematopoiesis regulating
CC	activity, tissue growth factor activity, immunomodulatory activity and
CC	activin/inhibin activity and may be useful in the diagnosis and/or
CC	treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC	inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC	(AAK55582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC	sequence listing were missing at the time of publication
XX	
SQ	Sequence 699 AA;
XX	
Query Match	11.3%; Score 121; DB 4; Length 699;
Best Local Similarity	27.3%; Pred. No. 0.13; Indels 64; Gaps 10
Matches	59; Conservative 23; Mismatches 70;
DQ	5 SYDSTSSDSLVRGIGFKRPHTVAWPWA-----YPVTSYPPLSQPDLPIRSPQ 56
Db	362 SSDDSSSDS-----SEDDKAPSKAGTGNSSNKPATTTSPAVKPAAP----KQ 408
DQ	57 PLGGSHRTPPSRRRSDGANSVAATYENEBPACEDADEDEBDYHNGYLVLDPDTPATSTA 116
Db	409 PVGGGQQLTRKADSSSSSEBSSESSEBKTK-----MVATTKPKATKA 453
DQ	117 APSAPALSTP-GIRDSAPSMESIDYVNPESGSASASLDGSRFYVVS--OELHPGAAX 174
Db	454 ALSLPQAQAQGGRHS--SSNS-----DSSSEBEETKSAYVKKKQKVAGGAAP 503
DQ	175 TEPAAL-----SQEAEEVEE-----GAP 194
Db	504 SKPASAKKGKAESNSSSDSSDSSSEBEKTLKGKGP 539

AA080140
ID AA080140 standard; protein; 718 AA.
XX
AC AA080140;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 3786.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; hematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US004098.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Dmanac RT, Aundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhtman T, Goodrich R;
XX
DR WPI: 2001-476283/51.
XX
DR N-PSDB; AAK53273.
XX
PT Nucleic acid encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
PS Claim 20; Page 435; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAW78323-AAW80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, hematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activity/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAW80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
XX
SQ Sequence 718 AA;
XX
Query Match 11.3%; Score 121; DB 4; Length 718;
Best Local Similarity 27.3%; Pred. No. 0.13; Indels 64; Gaps 10;
Matches 59; Conservative 23; Mismatches 70;
XX
QY 5 SYDSTSDSLYPRGIGKRPHTVAWPMPA-----YPVTSYPPISQDPLPPIPSPO 56
DB 381 SSSSDSDS-----SEDDAPSKRAGTTKSSNNKPAVTTKSPAVKPAAP-----KQ 427
QY 57 PLGSGHRTPSRRSDGANSVASYENEPACEDADEDEDYHNGYLVLPDSTPATSTA 116
DB 428 PVGGGKLLTRKADSSSEESSSEBEKTKK-----MVATTKPKATAXA 472

QY 117 APSAPALSTP-GIRDSAFMSIESIDYVNVPESGSASASIDGSRVYVNS-QELHPGAK 174
DB 473 ALSLPAKQAPQGRSDS--SSDS-----DSSSEBEKTKSKANVKKPKVAGGAAP 522
QY 175 TEPAL-----SSQAEVEVEE-----GAP 194
DB 523 SKPASAKGKAESSSSSDSDSSSEBEKTKGKGP 558
RESULT 13
ADP04241
ID ADP04241 standard; protein; 742 AA.
XX
AC ADP04241;
XX
DT 09-SEP-2004 (first entry)
XX
DE Human colon specific protein SEQ ID NO:222.
XX
KW human; colon specific nucleic acid; CSNA; colon specific protein; CSP;
KW cytostatic; vaccine; gene therapy; colon cancer.
XX
OS Homo sapiens.
XX
PN WO2004050858-A2.
XX
PD 17-JUN-2004.
XX
PF 04-DEC-2003; 2003WO-US038608.
XX
PR 04-DEC-2002; 2002US-0431333P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Macina RA, Turner LR, Sun Y, Rodriguez M, Burcham TS;
XX
DR WPI: 2004-480622/45.
XX
XX Novel colon specific protein derived from normal and neoplastic colon
PT cell, useful as vaccine in treating colon cancer and in identifying,
PT diagnosing, monitoring, staging, imaging colon cancer and non-cancerous
PT disease state in colon.
XX
PS Claim 12; SEQ ID NO 222; 655pp; English.
XX
XX The invention relates to a novel colon specific protein (CSP) (I), and
CC the nucleic acid encoding it. A CSP of the invention has cytostatic
CC activity, and may have use in a vaccine, and in gene therapy. The CSP is
CC useful for determining the presence of a colon specific protein in a
CC sample. The nucleic acid encoding the CSP is useful for determining the
CC presence of a colon specific nucleic acid (CSNA) in a sample. The CSP and
CC CSNA are useful for diagnosing or monitoring the presence and metastases
CC of colon cancer in a patient. The method of administering a composition
CC comprising a CSP or CSNA is useful for treating a patient with colon
CC cancer. The CSP and CSNA are useful as vaccine for treating colon cancer
CC and non-cancerous disease states in colon. The present sequence
XX represents a CSP of the invention.
XX
SQ Sequence 742 AA;
XX
Query Match 11.3%; Score 121; DB 8; Length 742;
Best Local Similarity 27.3%; Pred. No. 0.14; Indels 64; Gaps 10;
Matches 59; Conservative 23; Mismatches 70;
XX
QY 5 SYDSTSDSLYPRGIGKRPHTVAWPMPA-----YPVTSYPPISQDPLPPIPSPO 56
DB 405 SSSSDSDS-----SEDDAPSKRAGTTKSSNNKPAVTTKSPAVKPAAP-----KQ 451
QY 57 PLGSGHRTPSRRSDGANSVASYENEPACEDADEDEDYHNGYLVLPDSTPATSTA 116
DB 452 PVGGGKLLTRKADSSSEESSSEBEKTKK-----MVATTKPKATAXA 496
QY 117 APSAPALSTP-GIRDSAFMSIESIDYVNVPESGSASASIDGSRVYVNS-QELHPGAK 174

Db 1628 ----PCOTAVTAAAISSAGPVAVETSSRP 1652

Search completed: November 21, 2004, 14:16:53
Job time : 168 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 14:11:50 : Search time 39 Seconds
(without alignments)
345,194 Million cell updates/sec

Title: US-09-597-920b-4_COPY_31_233

Perfect score: 1071

Sequence: 1 RLPGSYDSTSSDLSLPKRGIO.....EAEVEEAGAPYENLOELN 203

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/6C.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/6D.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	116	10.8	2090	US-09-538-092-1081	Sequence 1081, Ap
2	112	10.5	1958	US-07-945-283-2	Sequence 2, Appl
3	111	10.4	455	US-09-270-767-45531	Sequence 45531, A
4	110	10.3	571	US-09-252-991A-30533	Sequence 30533, A
5	106.5	9.9	1298	US-08-690-473-2	Sequence 2, Appl
6	106.5	9.9	1298	US-09-259-821A-2	Sequence 2, Appl
7	106.5	9.9	1298	US-08-843-659-2	Sequence 2, Appl
8	106.5	9.9	1298	US-08-825-288A-2	Sequence 2, Appl
9	105.5	9.9	4019	US-09-854-133-425	Sequence 425, App
10	104.5	9.8	802	US-09-823-240A-2	Sequence 2, Appl
11	104.5	9.8	1219	US-09-344-624-4	Sequence 4, Appl
12	104	9.7	2441	US-08-194-468-2	Sequence 2, Appl
13	104	9.7	2441	US-08-961-739-2	Sequence 2, Appl
14	104	9.7	2441	US-09-514-247A-8	Sequence 8, Appl
15	104	9.7	2441	US-09-686-316-2	Sequence 2, Appl
16	104	9.7	2442	US-09-514-247A-10	Sequence 10, Appl
17	104	9.7	2442	US-09-538-092-1170	Sequence 1370, Ap
18	101.5	9.5	315	US-09-270-767-46043	Sequence 46043, A
19	101	9.4	2414	US-08-227-536-2	Sequence 2, Appl
20	101	9.4	2414	US-08-227-536-2	Sequence 2, Appl
21	101	9.4	2414	US-09-538-092-1189	Sequence 1289, Ap
22	101	9.4	2414	PCT-US95-04682-2	Sequence 2, Appl
23	100.5	9.4	580	US-09-270-767-41648	Sequence 41648, A
24	100.5	9.4	961	US-09-538-092-1231	Sequence 1231, Ap
25	100.5	9.4	1187	US-08-320-559-28	Sequence 28, Appl
26	100.5	9.4	1187	US-08-545-860D-28	Sequence 28, Appl
27	100.5	9.4	1187	PCT-US94-04496-28	Sequence 28, Appl

28	100.5	9.4	1210	1	US-08-320-559-26	Sequence 26, Appl
29	100.5	9.4	1210	3	US-08-545-860D-26	Sequence 26, Appl
30	100.5	9.4	1210	4	US-09-538-092-1179	Sequence 1179, Ap
31	100.5	9.4	1210	5	PCT-US94-04496-26	Sequence 26, Appl
32	98.5	9.2	174	3	US-09-199-637A-63	Sequence 63, Appl
33	98.5	9.2	335	2	US-08-405-175A-6	Sequence 6, Appl
34	98.5	9.2	346	3	US-09-147-236-11	Sequence 11, Appl
35	98.5	9.2	344	4	US-09-522-474-11	Sequence 11, Appl
36	98	9.2	750	3	US-09-165-239A-4	Sequence 4, Appl
37	98	9.2	943	2	US-08-469-537A-107	Sequence 107, App
38	98	9.2	1560	4	US-09-264-512B-2	Sequence 2, Appl
39	97.5	9.1	322	4	US-09-248-796A-17229	Sequence 17229, A
40	97.5	9.1	706	4	US-09-538-092-957	Sequence 957, App
41	97.5	9.1	8991	4	US-08-714-741-32	Sequence 32, Appl
42	97	9.1	282	1	US-07-712-476A-5	Sequence 5, Appl
43	97	9.1	1317	3	US-09-083-521-7	Sequence 7, Appl
44	97	9.1	1970	4	US-09-538-092-1005	Sequence 1005, Ap
45	96.5	9.0	553	3	US-09-083-351-2	Sequence 2, Appl

ALIGNMENTS

```
RESULT 1
US-09-538-092-1081
; Sequence 1081, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15566-542
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratSeqFormatter Version 0.9
; SEQ ID NO 1081
; LENGTH: 2090
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P35658
US-09-538-092-1081
Query Match 10.8%; Score 116; DB 4; Length 2090;
Best Local Similarity 24.9%; Pred. No. 0.18;
Matches 52; Conservative 20; Mismatches 85; Indels 52; Gaps 7;
QY 5 SYDSSSSSLYPRGQFGRPHTVAMPAPVPTVSYPPLSQDPLPIRSPQ-PLGSGHR 63
DB 1440 SFGSQQNTSTYV-----PSAPPPTVATATLPISFPLSGSLSSATPPLPWSAGRS 1492
QY 64 T-----PSRRDSDGANSVASENEPACEDADEDDYHNGYLVVVPDSTPATST 115
DB 1493 TBEATSSALPEKPGDSEVSASASLLEEQSANO-----LPQAPPTSD 1535
QY 116 AAPSAFALSTGIRDS-----AFSMESIDYVNVPSGSEASLDGSERYNVS 165
DB 1536 SVKEPVLVQPAVNSGTPAASSTSLVALSAGATATTCVPA--RTEAVPPASSFSV--- 1590
QY 166 QELHPGAKTPPALSSQEAEEVEEGAP 194
DB 1591 ----PGQAVTAAAISSAGPVAVETSTP 1615
RESULT 2
US-07-945-283-2
```

```

; Sequence 2, Application US/07945283
; Patent No. 5352596
; GENERAL INFORMATION:
; APPLICANT: Cheung, Andrew K.
; TITLE OF INVENTION: Pseudorabies Virus Deletion Mutants
; TITLE OF INVENTION: Involving The BPO and LIT Genes
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis P. Ribando
; STREET: 1615 No. 5352596th University Street
; CITY: Peoria
; STATE: IL
; COUNTRY: USA
; ZIP: 61604
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/945,283
; FILING DATE: 19920911
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Ribando, Curtis P
; REGISTRATION NUMBER: 27976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 309-685-4011 ext. 513
; TELEFAX: 309-685-4128
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1958 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-945-283-2

Query Match          10.5%; Score 112; DB 1; Length 1958;
Best Local Similarity 35.3%; Pred. No. 0.37;
Matches 30; Conservative 5; Mismatches 26; Indels 24; Gaps 3;

QY      29 PMPAYPVPTYSYPLSQPDLPIPRSPQLGGS-----HRTSSRRDS 71
Db      483 PSEPPRP-----PLPPPPPPPPPPPPAGSARRRRGGCGGPGGRRRGKRRRA 538
QY      72 DGANSVASYENEPACEDADEDD 96
Db      539 EGTEAAADAEE---EDGDEDEDE 560

RESULT 3
US-09-270-767-45531
; Sequence 45531, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 45531
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; US-09-270-767-45531

Query Match          10.4%; Score 111; DB 4; Length 455;
Best Local Similarity 21.3%; Pred. No. 0.053;
Matches 57; Conservative 23; Mismatches 106; Indels 82; Gaps 9;

```

```

QY      4 GSYDSTSSDLYRGIOFKRP-----HTVAPWPPAY---PVT----- 39
Db      165 GSNNTWMSFFIFLHFNTPTRQQQPRQGNVLPAHQPTPPFGSAPPAVSSNNFS 224
QY      40 -----YPLSQPDLPIPRSPQP-----LGSHRTSSRRDS-- 71
Db      225 GQTPMFAPIINHHPAPMGAPVVLISIPSPMPASLPWNSPLFKIPPLQAPAKSNGNQ 284
QY      72 -----DGANSVASYENEPACEDA-----DEDDDYHNPGLVVLPGSTPA 112
Db      285 NDVDVNCNPFSTISQESQAVANASAMPSCVPHGPADASDQDDDDMD--LVQDDDED 342
QY      113 TSTAAPSPALSTPGIRDSAFMESIDY-----VNPESGSAEASLDGSRBYVNSQ 166
Db      343 TDIPPLPGP---EBEPVEKVPKSSDDDLVEPENPTEPEPEPMFEESCDAVTEKSESS 399
QY      167 ELHPGAKTPEPALSSQEAEEVEEGAP 194
Db      400 DHEPSNSNVQAAAVENTDAEEARTSTP 427

RESULT 4
US-09-252-991A-30533
; Sequence 30533, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; CURRENT APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30533
; LENGTH: 571
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-30533

Query Match          10.3%; Score 110; DB 4; Length 571;
Best Local Similarity 22.1%; Pred. No. 0.11;
Matches 54; Conservative 30; Mismatches 92; Indels 68; Gaps 9;

```

```

QY      9 TSSDSLVRGIGFKRPHTVAPWPPAY-----PVTSPPLSQPDLPIPRSPQ 56
Db      315 TPTPTVSPSG-----SVARQAPAVSARVASTOAREPASVAPVDEBPLVVSHPQ 367
QY      57 PLGSHRTPS-----SRSDGANSV-ASYENEPACEDADE--DDYHNPGLVVLPD 108
Db      368 IAGKTHRPQPGGFPKTAIEVAISTQASVQSPPAFTVGTGSGRGERQPG--ETDPS 425
QY      109 STPATSTAPSPALSTPGIR-----DSAFMESIDYVNVPESESA 151
Db      426 ALPEDDAPVPLPMQPGGRILVRLASSGSRPLPLADLRLDVAQRIQVAAASH 485
QY      152 EASLDGSRBYVNSQELHPGAKTPEPALSSQEAEEVEEGAD-----YENL 199
Db      486 AARL-----QVRLPOLGAVEVQVHLGHQQLQVEISASPSGLAFLQAGELLERL 535
QY      200 QELN 203
Db      536 QRLH 539

RESULT 5
US-08-690-473-2
; Sequence 2, Application US/08690473
; Patent No. 5876923
; GENERAL INFORMATION:

```

```

; APPLICANT: Leopardi, Rosario
; APPLICANT: Roizman, Bernard
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 AS AN
; TITLE OF INVENTION: INHIBITOR OF APOPTOSIS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/690,473
; FILING DATE: 26-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: ARCD:239
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1298 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
US-08-690-473-2

Query Match          9.9% Score 106.5; DB 2; Length 1298;
Best Local Similarity 24.0%; Pred. No. 0.66;
Matches 49; Conservative 19; Mismatches 73; Indels 63; Gaps 7;

QY 52 PRSPPLGSHRTSPSSRRSDGANSVASYENEE-----PAC 87
DB 9 PGSPPTDGPPTTSPDRDERGALMGCA-ETEEGDDDPDHPDHDLDARDDRARA 67
QY 88 -EDADEDDYHNPGLVYL---PDSTPAISTAPAPALSTPGIRDSAFMESIDYV 142
DB 68 GTDAGEDGDGAVSPQLALASMEVEAVRTTPTPDPAASPPRTPAFRADDDGDEYDPA 127
QY 143 N-----VESGESASASIDGSRVYVNGQLHP-----G 171
DB 128 DAAGRAARARGRERAPLRGA--YDPPTDRLSPRPPAPRRRRHGRWRPSASTSSDSG 185
QY 172 AAKTEPALSSQAEVEVEEGAPD 195
DB 186 SSSSSASASSSSSDDEDDEDDGND 209

RESULT 6
US-09-259-821A-2
; Sequence 2, Application US/09259821A
; Patent No. 6210926
; GENERAL INFORMATION:
; APPLICANT: LEOPARDI, ROSARIO
; APPLICANT: ROIZMAN, BERNARD
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 IS AN INHIBITOR OF APOPTOSIS
; FILE REFERENCE: ARCD:317
; CURRENT APPLICATION NUMBER: US/09/259,821A
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 08/690,473
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
```

```

; LENGTH: 1298
; TYPE: PRT
; ORGANISM: HERPES VIRUS, TYPE 1
;
US-09-259-821A-2

Query Match          9.9% Score 106.5; DB 3; Length 1298;
Best Local Similarity 24.0%; Pred. No. 0.66;
Matches 49; Conservative 19; Mismatches 73; Indels 63; Gaps 7;

QY 52 PRSPPLGSHRTSPSSRRSDGANSVASYENEE-----PAC 87
DB 9 PGSPPTDGPPTTSPDRDERGALMGCA-ETEEGDDDPDHPDHDLDARDDRARA 67
QY 88 -EDADEDDYHNPGLVYL---PDSTPAISTAPAPALSTPGIRDSAFMESIDYV 142
DB 68 GTDAGEDGDGAVSPQLALASMEVEAVRTTPTPDPAASPPRTPAFRADDDGDEYDPA 127
QY 143 N-----VESGESASASIDGSRVYVNGQLHP-----G 171
DB 128 DAAGRAARARGRERAPLRGA--YDPPTDRLSPRPPAPRRRRHGRWRPSASTSSDSG 185
QY 172 AAKTEPALSSQAEVEVEEGAPD 195
DB 186 SSSSSASASSSSSDDEDDEDDGND 209

RESULT 7
US-08-843-659-2
; Sequence 2, Application US/08843659
; Patent No. 6218103
; GENERAL INFORMATION:
; APPLICANT: Leopardi, Rosario
; APPLICANT: Roizman, Bernard
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS US3 AND ICP4 AS
; TITLE OF INVENTION: INHIBITORS OF APOPTOSIS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/843,659
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: ARSB:519
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1298 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
US-08-843-659-2

Query Match          9.9% Score 106.5; DB 3; Length 1298;
Best Local Similarity 24.0%; Pred. No. 0.66;
Matches 49; Conservative 19; Mismatches 73; Indels 63; Gaps 7;

QY 52 PRSPPLGSHRTSPSSRRSDGANSVASYENEE-----PAC 87
```

Db 9 PGSPGPTGPPPTSPDRDERGALWGGA-ETEEGGDDPDHDPHDLDDARRGRAPAA 67
Qy 88 -EDADEDEDDYHNGYLVV----PDSTPATSTAPASAPALSTGIDSAFMSIDYV 142
Db 68 GTDAGEDAGDVSPROLALLASWVEAVRTIPTDPAPASPPRTAPFRADDDGDYDAA 127
Qy 143 N-----VPEGESAEASLDGSEYVNVSQLHP-----G 171
Db 128 DAAGDRAPAREREAPLRGA--YDPPTDRLSPRPAPAPRRRRHGRWRPESASTSSDSG 185
Qy 172 AAKTEPALSSQAEVEVEEGAPD 195
Db 186 SSSSSSSASSSSSSDEDEDDGND 209

RESULT 8
US-09-825-288A-2
Sequence 2, Application US/09825288A
Patent No. 6723511
GENERAL INFORMATION:
APPLICANT: LEOPARDI, ROSARIO
APPLICANT: ROIZMAN, BERNARD
TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICPI IS AN INHIBITOR OF APOPTOSIS
FILE REFERENCE: ABCD.3170SC1
CURRENT APPLICATION NUMBER: US/09/825,288A
CURRENT FILING DATE: 2001-04-02
PRIOR APPLICATION NUMBER: 09/4259,821
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: 08/690,473
PRIOR FILING DATE: 1996-07-26
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 1298
TYPE: PRT
ORGANISM: HERPES VIRUS, TYPE 1
US-09-825-288A-2

Query Match 9.9%; Score 106.5; DB 4; Length 1298;
Best Local Similarity 24.0%; Pred. No. 0.66; Indels 63; Gaps 7;
Matches 49; Conservative 19; Mismatches 73;
Qy 52 PRSPQPLGSHRTSPSRSDGANSVASYENEE-----PAC 87
Db 9 PGSGPTGDPPTSPDRDERGALWGGA-ETEEGGDDPDHDPHDLDDARRGRAPAA 67
Qy 88 -EDADEDEDDYHNGYLVV---PDSTPATSTAPASAPALSTGIDSAFMSIDYV 142
Db 68 GTDAGEDAGDVSPROLALLASWVEAVRTIPTDPAPASPPRTAPFRADDDGDYDAA 127
Qy 143 N-----VPEGESAEASLDGSEYVNVSQLHP-----G 171
Db 128 DAAGDRAPAREREAPLRGA--YDPPTDRLSPRPAPAPRRRRHGRWRPESASTSSDSG 185
Qy 172 AAKTEPALSSQAEVEVEEGAPD 195
Db 186 SSSSSSSASSSSSSDEDEDDGND 209

RESULT 9
US-09-854-133-425
Sequence 425, Application US/09854133
Patent No. 6759508
GENERAL INFORMATION:
APPLICANT: Lodes, Michael J.
APPLICANT: Mohamach, Radoch
APPLICANT: Henderson, Robert A.
APPLICANT: Benson, Darin R.
APPLICANT: Secrist, Heather
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
FILE REFERENCE: 210121.475C10
CURRENT APPLICATION NUMBER: US/09/854,133

CURRENT FILING DATE: 2001-05-11
NUMBER OF SEQ ID NOS: 735
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 425
LENGTH: 4019
TYPE: PRT
ORGANISM: Homo sapiens
US-09-854-133-425

Query Match 9.9%; Score 105.5; DB 4; Length 4019;
Best Local Similarity 27.0%; Pred. No. 3.8; Indels 49; Gaps 11;
Matches 61; Conservative 19; Mismatches 97;
Qy 3 GSYDSTSSDGLYPRGIGFRRPHT-----VAPV-----PPAYPVVTSYP-----PLSQP 46
Db 1218 PGTRPRTTVDP-YSQDQGTREPSTQTDLPVTYTNQHSRDPVHAPGTRPGISVPSQ 1276
Qy 47 DLLPPIRSPQPLGSHRT-PSRRSDGANSVASYENEEPCADDEDDYHNGYLVV 105
Db 1277 PATPRPRISEGFRRSMTFRVLPNODPFLQAA--QNRGPAI-----PGPLVR 1322
Qy 106 LPSTPATSTAPASAPALSTGIDSAFMSIESIDYVNVPEGESAEASLDGSEYVNV 165
Db 1323 PPPT--CSQTPRPFGSLDTSFRVSPSAR--DPYDQSMTRRSQSDSGTGTQTAHVA 1378
Qy 166 QELHPG-----AAKTEPALSSQAEVEVE-----EGAPDYEN 198
Db 1379 DQRRPGESEGFCASSNSPMHSGQGFSGVSQLGCPVTSVTTQTN 1424

RESULT 10
US-09-823-240A-2
Sequence 2, Application US/09823240A
Patent No. 6716597
GENERAL INFORMATION:
APPLICANT: Frank B. Gertler
APPLICANT: James E. Bear
APPLICANT: Jurgen Lourelo
APPLICANT: Joseph Lourelo
TITLE OF INVENTION: Methods and Products for Regulating Cell
FILE REFERENCE: M00656.70064.US
CURRENT APPLICATION NUMBER: US/09/823,240A
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 60/194,564
PRIOR FILING DATE: 2000-04-03
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 802
TYPE: PRT
ORGANISM: Mus musculus
US-09-823-240A-2

Query Match 9.8%; Score 104.5; DB 4; Length 802;
Best Local Similarity 24.0%; Pred. No. 0.52; Indels 77; Gaps 9;
Matches 60; Conservative 17; Mismatches 96;
Qy 4 GSYDSTSSDGLYPRGIGFRRPHTVAPVPPVTSYPPLSQDLP-----PRSP 55
Db 422 GSIDSVT---YVSPPTSGPAAPPPPPPPPPPPPPPLPLPLPLSLHSGSQASP 477
Qy 56 QPLGSHRTSPSRD---SDGANSVASYEN-----ESPACEDDEDDDD-----96
Db 478 PGTPLASTPSKPSVLPSPSAGAPASAEPLPLPELIDSSASBPGLQAAISQPAESPPOG 537
Qy 97 -----YHNGYLVV-----PDSTPATSTA-----116
Db 538 LVLGPPAPPPPPPLPSGAPASALPPPPGPPPPPLPSTGPPPPPPPPPPPLPQAPPPPP 597
Qy 117 APSAPALSTGIDSAFMSIESIDYVNVPEGESAEASLDGSEYVNVSQLHPGAATE 176
Db 598 PPAPPLPAGI-----FSGSTSD--NRPLTGAAAIAGAKLRKVSHEDESGFPGGNTG 651

QY 177 PAALSSOEA 186
DB 652 SVSLASSKAD 661

RESULT 11

US-09-344-624-4
; Sequence 4, Application US/09344624
; Patent No. 6753154
; GENERAL INFORMATION:
; APPLICANT: Chen, Hui-Mei
; APPLICANT: Bissell, Mina
; TITLE OF INVENTION: HUMAN AZ-1 GENE, VARIANTS THEREOF AND EXPRESSED GENE
; TITLE OF INVENTION: PRODUCTS
; FILE REFERENCE: 2960.44 (HV)
; CURRENT APPLICATION NUMBER: US/09/344,624
; CURRENT FILING DATE: 1999-06-25
; EARLIER APPLICATION NUMBER: 60/090,747
; EARLIER FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1219
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-344-624-4

Query Match 9.8%; Score 104.5; DB 4; Length 1219;
Best Local Similarity 22.5%; Pred. No. 0.92;
Matches 64; Conservative 28; Mismatches 73; Indels 119; Gaps 16;

QY 5 SYDS-----TSSDSLYPRGICQ---FKRPHYAP--WPPAYP-----VTSYPL 43
DB 280 SYHSDVQGVSTDLIAQSSDSEBAFETPESTTPVKAPAPAPPPPEPEVPEPEVSTOPPP 339
QY 44 SOP-----DLPLI---PRSPQPLGSHRTPPS-----RRSDGANSYA 78
DB 340 EEPGCGSTVPVPPGPRSDSVEGSPFRPPSHFSAVPDEDOPIASSGYNLDFDIEILD 399
QY 79 SYNEBPACEDADEDE-----DDYH-----NPGYLV 104
DB 400 TFOGLEPASPADKONGKVRNRKSTDSVPIKSTLSLSQASDFGASSGNGPAVA 459
QY 105 VLFP-----STPATS-----TAAAPAL-----STGIRDSARSMESIDYVNP 145
DB 460 LAPPAVSTGSSSSASTLKRTKKRPPPSLKQKQTKKPTETPVKE---TQGEPPDESLVP 516
QY 146 ESGESAASLDGSRVYVNSOELHPGAKTE---PAALSSOEA 186
DB 517 SGE-----MLASSTKTESAKTBGSPALLETPE 546

RESULT 12

US-08-194-468-2
; Sequence 2, Application US/08194468
; Patent No. 5750336
; GENERAL INFORMATION:
; APPLICANT: Montminy, Marc R.
; TITLE OF INVENTION: ASSAYS FOR THE IDENTIFICATION OF
; TITLE OF INVENTION: COMPOUNDS WHICH INHIBIT ACTIVATION OF CAMP AND MITOGEN
; TITLE OF INVENTION: RESPONSIVE GENES
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESS: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/194,468
; FILING DATE: 10-FEB-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter, Stephen E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: P41 9672
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)-546-4737
; TELEFAX: (619)-546-9392
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2441 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-194-468-2

Query Match 9.7%; Score 104; DB 1; Length 2441;
Best Local Similarity 25.8%; Pred. No. 2.6;
Matches 54; Conservative 18; Mismatches 79; Indels 58; Gaps 10;

QY 3 PGSDSTSSDLYPRGICQFKRPHYAPWP--PAYPPVTSYPLGQDLPPIPRSPQPIGG 60
DB 904 PGSVPSAAQTOSTPT-VQAAQAQVTPQPTVPQPPSVATQSSQQ--PPVHTQPPG- 959
QY 61 SHRPSSRRSDGANSVSYNEBPACEDADEDEDYNNPYLVLPSTPATSTAAPSA 120
DB 960 ---TPLSQ-----AAASINRVF-----TSTYTSATS 985
QY 121 PALSTPGIRDSAFSMESIDYVNVPSGES-----AEASLDGSRVYVNSOELHPGA 172
DB 986 SQQCPQDVPMLEMTETVQTDAB-PEPTESKGEPRSEKMEEDLQSSQ--VKERTDTTE 1041
QY 173 AKTEPALSSQ-----EAEVESEGAPD 195
DB 1042 QKSEPEVEERKPEYKVEAKEEENSSND 1070

RESULT 13

US-08-961-739-2
; Sequence 2, Application US/08961739A
; Patent No. 6063583
; GENERAL INFORMATION:
; APPLICANT: Montminy, Marc R.
; TITLE OF INVENTION: Methods for Treating Diabetes Mellitus
; FILE REFERENCE: SALK1650-1
; CURRENT APPLICATION NUMBER: US/08/961,739A
; CURRENT FILING DATE: 1997-10-31
; EARLIER APPLICATION NUMBER: US 194,468
; EARLIER FILING DATE: 1994-02-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2441
; TYPE: PRT
; ORGANISM: Mus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(2441)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-08-961-739-2

Query Match 9.7%; Score 104; DB 3; Length 2441;
Best Local Similarity 25.8%; Pred. No. 2.6;
Matches 54; Conservative 18; Mismatches 79; Indels 58; Gaps 10;

QY 3 PGSDSTSSDLYPRGICQFKRPHYAPWP--PAYPPVTSYPLGQDLPPIPRSPQPIGG 60
DB 904 PGSVPSAAQTOSTPT-VQAAQAQVTPQPTVPQPPSVATQSSQQ--PPVHTQPPG- 959

```
QY 61 SHRTSSNRSDSGANSVANSYENEEBACADADEDEDYNNPGYLVLPSTPATSTAAASA 120
    |||  |||
Db 960 ---TFLSQ-----AAASIDNRVD-----TSTVTSATST 985
    |||  |||
QY 121 PALSTPGIRDSAFSMESIDYVNVBESGES-----AAASLDGSRVYVNSOELHPCA 172
    |||  |||
Db 986 SQQSPEDVPMLEKTEVOTDDAE-DEPTESKKEPRSEMMEDLQSSQ---VKETDTTE 1041
    |||  |||
QY 173 AKTEPALSSQ-----EAEVEEBGAPD 195
    |||  |||
Db 1042 QKSEPMVEYBEKKPEYVKAKEEENSSND 1070
```

```

RESULT 14
US-09-514-247A-8
: Sequence 8, Application US/09514247A
: Patent No. 6365361
: GENERAL INFORMATION:
: APPLICANT: TANABE SEIYAKU CO. LTD.
: APPLICANT: TANIGUCHI, Tomoyasu
: APPLICANT: MIZUKAMI, Junko
: TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND ANTAGONIST TO PPR
: FILE REFERENCE: TANIGUCHI-6
: CURRENT APPLICATION NUMBER: US/09/514,247A
: CURRENT FILING DATE: 2000-02-28
: PRIOR APPLICATION NUMBER: PCT/JF98/03734
: PRIOR FILING DATE: 1998-08-24
: PRIOR APPLICATION NUMBER: JP231084/1997
: PRIOR FILING DATE: 1997-08-27
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO: 8
: LENGTH: 2441
: TYPE: PRT
: ORGANISM: mouse
US-09-514-247A-8

```

Query Match	9.7%	Score 104	DB 3	Length 2441
Similarity	25.8%	Pred. NO. 2.6		
Best Local Match	54	Conservative 18	Mismatches 79	Indels 58
				Gaps 10
Qy	3	PGSYDSTSSDLSLYPRGIQFKRHVTAAMP--PAYPVVTSYTPSLQPDLLPIRSDQPIYG	60	
Db	904	PGSVPSAAQOTGSTR-VQAAAGVATPGQGTVPQGFVATPGSSQGC--PFRVHTQPG	959	
Qy	61	SHRTSSRRDDGANSVASYENEPACEDADEDDYNNPGVLVLPSTPATSTAAASA	120	
Db	960	--TFLSQ-----AAASIDNRVP-----TSTVTSASTS	985	
Qy	121	PALSTPGIRDSAFSMESIDYVNVPESGES-----AASLDGSRVYVNSQELHGA	172	
Db	986	SQGPSPVPMLEMKTEVQTDAAE-PEPESKKEPRSEMMEDLQSSGC--VKERTDTTE	1041	
Qy	173	AKTEPALSSQ-----EAEVEEGCAPD	195	
Db	1042	QKSEMEVEEKKPEYKVAKEEENSSND	1070	

RESULT 15
US-09-686-316-2
Sequence 2, Application US/09686316
Patent No. 6646115
GENERAL INFORMATION:
APPLICANT: MONTMAY, Marc R.
TITLE OF INVENTION: Methods for Treating Diabetes Mellitus
FILE REFERENCE: SAL1650-1
CURRENT APPLICATION NUMBER: US/09/686,316
CURRENT FILING DATE: 2000-10-10
PRIOR APPLICATION NUMBER: US/08/961,739
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: US 194,468
PRIOR FILING DATE: 1994-02-10

```

? NUMBER OF SEQ ID NOS: 4
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 2
? LENGTH: 2441
? TYPE: PRT
? ORGANISM: Mus
? FEATURE:
? NAME/KEY: VARIANT
? LOCATION: (1)...(2441)
? OTHER INFORMATION: Xaa = Any Amino Acid
US-09-686-316-2

```

Query Match	Score	DB	Length
9.7%	104	4	2441

Best Local Similarity 25.8%; Pred.NO. 2.6;
Matches 54; Conservative 18; Mismatches 79; Indels 58; Gaps 10;

```

Qy 3 PGVSDFSTSDILYVRGQIFKRPHVIAWMP--FAVPEVTSYPTLSQPDLLPIPRSBQPGG 60
Db 504 PGSVPSAAQTOSTTT-VQAAAQAVTQPCQCPVUPPSPVATQSSQQQ--PIPVHTQPGG- 959
Qy 61 SHRTPESSRRSDGANSVASYENEEBPACEDADEDEDYHNPGLVLLPDSPTPATSTAESA 120
Db 960 ---TFLSQ-----AAASIDNRVP-----TPSTVTSATS 985
Qy 121 PALSTPGIRDSAFSMESIDYVNVPSGES-----AASLDGSKREYVNVQGLHGA 172
Db 986 SQQDGPVPMLEMKTEVQTDAAE-PEFTESKGEPRSEMKMEEDLQSSSQ--VKETDTTE 1041
Qy 173 AKTEPALASQ-----EAEVEEGEAPD 195
Db 1042 QKSEPEVEEEKPEVKVAAKEEENSSND 1070

```

Search completed: November 21, 2004, 14:21:37
Job time : 41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 14:20:19 ; Search time 141 Seconds
(without alignments)
509.843 Million cell updates/sec

Title: US-09-597-920B-4_COPY_31_233

Perfect score: 1071

Sequence: 1 RLPGSYDSTSSDLVPRGIQ.....EAEVEBEGADYENLQELIN 203

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1570615 seqs, 354127592 residues

Total number of hits satisfying chosen parameters: 1570615

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :	Published Applications AA:*
1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9:	/cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10:	/cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13:	/cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16:	/cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17:	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
18:	/cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
19:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
20:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	123.5	11.5	311	US-10-437-963-174234	Sequence 174234,
2	120.5	11.3	468	US-10-437-963-146700	Sequence 146700,
3	114.5	10.7	1000	US-10-128-714-3305	Sequence 3305, Ap
4	114	10.6	762	US-10-437-963-131253	Sequence 131253,
5	113.5	10.6	735	US-10-282-122A-69392	Sequence 69392, A
6	112.5	10.5	1177	US-10-193-692-4	Sequence 4, Appli
7	112.5	10.5	1186	US-10-193-692-2	Sequence 2, Appli
8	112	10.5	430	US-10-739-930-5638	Sequence 9638, Ap
9	112	10.5	519	US-10-113-794A-2	Sequence 2, Appli
10	112	10.5	519	US-10-428-487-14	Sequence 14, Appli
11	112	10.5	519	US-10-258-371B-28	Sequence 28, Appli
12	112	10.5	923	US-10-114-270-152	Sequence 152, App
13	111	10.4	169	US-10-425-115-224853	Sequence 224853,

14	110	10.3	346	US-10-437-963-184670	Sequence 184670,
15	110	10.3	584	US-10-156-761-12405	Sequence 12405, A
16	109.5	10.2	340	US-10-437-963-181586	Sequence 181586,
17	109.5	10.2	1151	US-09-825-751A-79	Sequence 79, Appli
18	109.5	10.2	1240	US-10-369-493-4031	Sequence 4031, Ap
19	109	10.2	1678	US-10-437-963-138217	Sequence 138217,
20	108.5	10.1	1343	US-10-408-765A-1085	Sequence 1085, Ap
21	108.5	10.1	1714	US-10-128-714-3176	Sequence 3176, Ap
22	108.5	10.1	1750	US-10-128-714-8176	Sequence 8176, Ap
23	108	10.1	509	US-10-437-963-106493	Sequence 106493,
24	107.5	10.0	634	US-10-108-260A-3340	Sequence 3340, Ap
25	106.5	9.9	102	US-09-739-907-58	Sequence 58, Appli
26	106.5	9.9	102	US-09-938-671-58	Sequence 58, Appli
27	106.5	9.9	180	US-09-739-907-106	Sequence 106, App
28	106.5	9.9	180	US-09-938-671-106	Sequence 106, App
29	106.5	9.9	420	US-10-425-115-184837	Sequence 184837,
30	106.5	9.9	1298	US-09-825-288A-2	Sequence 2, Appli
31	106	9.9	1372	US-10-437-963-165419	Sequence 165419,
32	106	9.9	1790	US-10-437-963-138166	Sequence 138166,
33	105.5	9.9	283	US-09-864-761-34492	Sequence 34492, A
34	105.5	9.9	553	US-10-437-963-147077	Sequence 147077,
35	105.5	9.9	4019	US-09-738-973-425	Sequence 425, App
36	105.5	9.9	4019	US-09-854-133-425	Sequence 425, App
37	105.5	9.9	4019	US-10-144-649A-425	Sequence 425, App
38	105	9.8	1164	US-10-437-963-165343	Sequence 165343,
39	105	9.8	1345	US-10-437-963-165375	Sequence 165375,
40	104.5	9.8	497	US-09-764-864-1314	Sequence 1314, Ap
41	104.5	9.8	534	US-09-764-864-861	Sequence 861, Ap
42	104.5	9.8	678	US-10-437-963-165341	Sequence 165341,
43	104.5	9.8	802	US-09-823-240-2	Sequence 2, Appli
44	104.5	9.8	1059	US-10-425-115-222100	Sequence 222100,
45	104	9.7	553	US-10-771-931-11	Sequence 11, Appli

ALIGNMENTS

RESULT 1
US-10-437-963-174234
; Sequence 174234, Application US/10437963
; Publication No. US20040123343A1
GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 174234
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Oryza sativa
; NAME/KEY: unsure
; LOCATION: (1) .. (311)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_72195C.1.pep
US-10-437-963-174234

Query Match 11.5%; Score 123.5; DB 16; Length 311;
Best Local Similarity 23.1%; Pred. No. 0.15;
Matches 46; Conservative 30; Mismatches 84; Indels 39; Gaps 6;

QY 30 WPRAYPTVTSIPPLSQPD---LPIPRSPQPLGSHRTPSSRRSDGANSVASYENEDPA 86

1


```

FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_33336C.1.pap
US-10-437-963-131253

Query Match
Best Local Similarity 10.6%; Score 114; DB 16; Length 762;
Matches 52; Conservative 21; Mismatches 65; Indels 54; Gaps 10;

QY 29 PMPAPVPTSTYSP-----LSQPDLPPIPSPPQLGSHHTPSRRSDGANSVASYENE 84
DB 42 PPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 89
QY 85 PACEDADEDEDY-----HNPQYVLPDSTPATSTA-----APSAPALSTPGIR 129
DB 90 EAEIEEEDDEDHLDRRRLRHRRPQVPPVSSPAAGTPTVTPQQPPPPPPPPPPPPPPPP 149
QY 130 --DSAPSE--S1---DDYVNPESGESASASLDSGREYVNVSGELHPGAATPEPA 179
DB 150 TWDFPFMSDEGMASIAPDDEDEIQEPE-----DEKTVPASPPRPPPLSPPTPVAA 198
QY 180 LSSQAEVEEER 191
DB 199 PAPQ--EEPEEE 208

RESULT 5
US-10-282-122A-69392
Sequence 69392, Application US/10282122A
GENERAL INFORMATION:
APPLICANT: Wang, Lianghu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haeselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See file Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 69392
LENGTH: 735
TYPE: PRT
ORGANISM: Pseudomonas syringae

```

```

FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (445T..(445)
OTHER INFORMATION: X=any amino acid
US-10-282-122A-69392

Query Match
Best Local Similarity 10.6%; Score 113.5; DB 15; Length 735;
Matches 55; Conservative 22; Mismatches 80; Indels 49; Gaps 11;

QY 23 RPHVAPMPAPVPTSTYSPPLSQPDLPPIPSPPQ-----LGSHTPTSSRRSDG---- 73
DB 413 QPPVAPAPAPVAPQPAKAP--APQIKEP--EPQXTQAAGRRRASAVERVESAAGRKA 469
QY 74 -----ANSVASYENE-----EPACEDADEDEDYHNPQYVLPDSTPATSTA 116
DB 470 CRARARARARASCAVEAPQEPVAPVLETVEQPD-----LTPMAPAPAPSP 520
QY 117 APSAP-ALSTPGIRSAFS--MESIDD--YVNPESGESASASLDSGREYVNVSGELHP 170
DB 521 VPDAPQAPSPVPEEQVTPAMLEAI PDSATVUSAPMDXDDPEPAD--DIVEPDIDIF 577
QY 171 GAAK--TEPPAALSQAEVEEER 194
DB 578 ASYSTLDELAHESVVELEAVEPEPAP 603

RESULT 6
US-10-193-692-4
Sequence 4, Application US/10193692
Publication No. US20030074682A1
GENERAL INFORMATION:
APPLICANT: Exelixis Deutschland GmbH
TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF A NOVEL TELBOST POTASSIUM
FILE REFERENCE: ARO2-005C
CURRENT APPLICATION NUMBER: US/10/193,692
CURRENT FILING DATE: 2002-07-11
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 1177
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION:
US-10-193-692-4

Query Match
Best Local Similarity 10.5%; Score 112.5; DB 14; Length 1177;
Matches 49; Conservative 20; Mismatches 61; Indels 25; Gaps 10;

QY 27 VAPMPAPVPTSTYSPPLSQPDLPPIPSPPQ--GSHRTPTSSRRSDGANSVASYENE 84
DB 1039 MAPVPAPVAV-SPDPLAHF--VP-PAHPTSLYTAANTTPSLQITD-----ASSPGKS 1089
QY 85 PACEDADEDEDYHNPQ-YVYVLPDST-----PATSTAPSA-PALSTGIRDSATSMESI 138
DB 1090 PDVDSLKERSPDLSGHLTVASTDTMSKSETELVSAGPILQPPGLCSSLRPPL 1149
QY 139 DDYVNPESGESASASLDSGREYV-NVSGELHPGA 172
DB 1150 PDSLGP-----GTLEGSPEIQAHVSDPVLPSS 1177

RESULT 7
US-10-193-692-2
Sequence 2, Application US/10193692
Publication No. US20030074682A1
GENERAL INFORMATION:
APPLICANT: Exelixis Deutschland GmbH
TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF A NOVEL TELBOST POTASSIUM
FILE REFERENCE: ARO2-005C
CURRENT APPLICATION NUMBER: US/10/193,692

```

```

; CURRENT FILING DATE: 2002-07-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1186
; TYPE: PRT
; ORGANISM: Danio rerio
US-10-193-632-2

```

Query Match 10.5%; Score 112.5; DB 14; Length 1186;
 Best Local Similarity 31.6%; Pred. No. 5.4;
 Matches 49; Conservative 20; Mismatches 61; Indels 25; Gaps 10;

Qy 27 VAPMPRAYEVTSTVPLSOPDLPLPRSPQL--GSHRRPSSRRSDDGANSVVSAYENE 84
Db 1048 MAPRPAYASAV-SEDPDAH--VP-PAHPSPSLYTTAAHNTTSPLOITD-----ASSPGKS 1098

Qy	85	PACDADDEDDYHNPQ-YVWLPST----	PATSTAAPA-PALSTPGRDSAFSMESI	138
Db	1099	PDVDSLKESPDLSLSGIRHTVASTDDTMS	SPETLSVEBAGPLLPGLCCSSLRFL	1158

QY 139 DDYNNVPEGGEAEEASLDGSRREYV-NVSGDELHPGA 172
| : | : | | : | : | : | :
Db 1159 PDSLEGP-----GTLESGPELGRHVSPDVLPGS 1186

RESULT 8
US-10-739-930-9638

Publication No. US20040216190A1
GENERAL INFORMATION:
APPLICANT: KOVALIC, MICHAEL K. ACTN MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
TITLE OR INVENTION: DIVERGIC ACTN MOLECULES AND OTHER MOLECULES ASSOCIATED WITH

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

```

; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 9638
; LENGTH: 430
; TYPE: DPM

```

```

; ORGANISM: Triticum aestivum
;
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) (430)

```

```
; OTHER INFORMATION: unsure at all Xaa location  
; FEATURE:  
; OTHER INFORMATION: Clone ID: TRIAE-23APR03-C1  
ITS-10-728-020-0638
```

Query Match	10.5%	Score 112;	DB 17;	Length 430
Best Local Similarity	25.7%	Pred. No. 1.7;		
Matches	54;	Conservative	35;	Mismatches 75;
				Indels 5

Oy 21 FKRP-HTVAWMPAPVPVTSYRPSQDILPIPRS-----PQLPGSHKTP----- 65
23 FKGSNHYR/GSGRPRASTGSSRP-SNPNNMNNPDPVVKQKDDPAPR/GSPFETTPETAVTSSG 79

Oy 66 SSRR-DSDCGANSVASYENEPACEDADEDED--DYHNPGLVLVDSTPATSTAA----- 117
ob 80 SSSPPDANNCPTVC-----PNCCHPEESHAUVEHUHC--CIAGCGAPAHNVAIVWC 121

```
Oy -----PSA-----PALSTPGIRDSAFSWE-----SIDDY 141
      | : |
118 ---| : |
      | : |
Db 112 NPBPBALEVVUKKICNTIKRQBNDDVVRVVLGNPBIKEJANARFQCI STLPAWERTIGD 190
      | : |
```

142 UNVPESGE-----SAAEASLDGSGREYVAVVSOELHPGAKTEPAALSQAEAEVEEBCGA 193
 191 -----FGGELFAVMDETPTAPRLSGTPOAVYLLPSPHSPDL-----DTLAGSRSSKCRSGV 243

QY	194 PDYENLQEL 202
	: : : :
Db	244 DEOKRVNKT 252

RESULT 9
US-10-113-794A-2

Publication No. US20030022202A1
GENERAL INFORMATION:
APPLICANT: Flanagan et al.
TITLE OF INVENTION: A METHOD OF IDENTIFICATION OF A PROTEIN COMPLEX

TITLE OF INVENTION: CHEMOATTRACTION
FILE REFERENCE: 2535/106
CURRENT APPLICATION NUMBER: US/10/113,794A
CURRENT FILING DATE: 2002-04-01

NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.
SEQ ID NO 2
SEQUENCE: 510

TYPE: PRT
ORGANISM: Homo sapiens
S-10-113-794A-2

Query Match	10.5%	Score 112;	DB 14;	Length 519;
Best Local Similarity	24.3%	Pred. No. 2.2;		
Matches	51;	Conservative	22;	Mismatches 79;
				Indels 58;
				Gaps 11;

57 PSQEPISKDSATSEG-----SPGPDAPSPKDVPPCQGEPPAQLSPCQDLPAQG 10

```

56 QPLGSHRTPSRRDSOGANSVASYENEPACEDADE-----DDYHNPYLVL 10
      :|::||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
108 EPL-PHQDPLLTKDLPAIQE--SPTRLPPCDDLPSGVSLPAKALTIEDTMSGDLLAA 16

```

107 PDSTPASTAAAPALSTPGIR-DSAFMSMEIDYVNPESGSAEALDGSREYVNS 16
 164 TGDP-----AAE-RPAFVPEVFLDSTYSQKA-----GAQCGSGDEEDAEAA 20

```

166 QELHPGAKTEPALSSQEAEEVEEGAPD 19
      :: |      : || | : |
207 EEVEEG-----EEGGEDEDEDTS D 22

```

ESULT 10
S-10-428-487-14
Sequence 14 Application RIS/10428487

Publication No. US20040006780A1
GENERAL INFORMATION:
APPLICANT: RASTELLI, LUCA K.
APPLICANT: GERARD, HANS-DIETER

TITLE OF INVENTION: VEGF-MODULATED GENES AND METHODS EMPLOYING THEM
FILE REFERENCE: 09800080-0103
CURRENT APPLICATION NUMBER: US/10/428,487
CURRENT FILING DATE: 2003-05-02

PRIOR APPLICATION NUMBER: 09/815,153
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,201
PRIOR FILING DATE: 2000-03-22

NUMBER OF SEQ ID NOS: 84
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 14
LENGTH: 519

TYPE: PRT
ORGANISM: Homo sapien
S-10-428-487-14

Query Match	10.5%	Score 112;	DB 15;	Length 519;
Best Local Similarity	24.3%;	Pred. No. 2.2;		
Matches	51;	Conservative	22;	Mismatches 79;
				Indels 58;
				Gaps 11;

3 PGSYDSRSSDSLYPRGIQFKRHTVAPNPAPVAVTSLYPQLSP----DLLP--IPRSP 55
57 PSQEPFLSSKDSATSEG-----SPFGPDAPESKDVYPCQEEPPAQDLSPCQDLPAQG 10

```

APPLICANT: Zerhusen, Bryan D.
APPLICANT: Gorman, Linda
APPLICANT: Shenoy, Suresh G.
APPLICANT: Pena, Carol E.A.
APPLICANT: Smithson, Glenda
APPLICANT: Burgess, Catherine E.
APPLICANT: Gerlach, Valerie
APPLICANT: Padigaru, Muralidhara
APPLICANT: Shukets, Richard A.
APPLICANT: Gangolli, Esna A.
APPLICANT: Taupier Jr., Raymond J.
APPLICANT: Caeman, Stacie J.
APPLICANT: Ji, Weizhen
APPLICANT: Anderson, David W.
APPLICANT: Liette, Mario W.
APPLICANT: Rastelli, Luca
APPLICANT: Edinger, Shlomit R.
APPLICANT: Stone, David J.
APPLICANT: Macdougall, John R.
APPLICANT: Rothenberg, Mark E.
FILE REFERENCE: No. US20040030110A1e1 Proteins and Nucleic Acids Encoding Same
CURRENT APPLICATION NUMBER: US/10/114,270
CURRENT FILING DATE: 2002-11-27
PRIOR APPLICATION NUMBER: 60/281,086
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/281,136
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/281,863
PRIOR FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/281,906
PRIOR FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/282,020
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,930
PRIOR FILING DATE: 2001-04-10
PRIOR APPLICATION NUMBER: 60/282,934
PRIOR FILING DATE: 2001-04-10
PRIOR APPLICATION NUMBER: 60/283,512
PRIOR FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/283,710
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: 60/284,234
PRIOR FILING DATE: 2001-04-17
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 470
SEQ ID NO 152
LENGTH: 923
TYPE: PRT
ORGANISM: Homo sapiens
US-10-114-270-152

Query Match          10.5%; Score 112; DB 15; Length 923;
Best Local Similarity 24.3%; Pred. No. 4,4; 79; Indels 58; Gaps 11;
Matches 51; Conservative 22; Mismatches

QY 3 PGSYDTSRSDSLPRGIQFRPHTVAPWPAPYPPVTSYPLSQP-----DLLP---IPRSP 55
Db 462 PQGEPLSKSKSATSSEG-----SPRGDAPPSKDVPPCCGPPPAQDUSPQQDLPAQG 512
QY 56 QPLGSGHRTSSRRSDSGANSVASYENEBACEDADE-----DDYHNPGLVVL 106
Db 513 EPL--PHQDPLTKDLPALIGE--SPTRDLPPCQDLPPSQVSLPAKALTEDTSSGDLAA 568
QY 107 PDSTPATSTAPAPALSTGIR-DSAFSMESIDDYVNNPESGESAPASLDSGREYVNN 165
Db 569 TGDPP-----AAP-RPAFVPIPEVRLDSTYSOKA-----GAEGCGSGDEDAEFA 611
QY 166 QELHPGAKTEPALSSQEAEEVEEGAPD 195
Db 612 EEVEEG-----EEGEEDEDEDTSD 630

```

RESULT 13
US-10-425-115-224853
; Sequence 224853, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/425.115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 224853
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(169)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_136658C.1.pep
US-10-425-115-224853

Query Match 10.4%; Score 111; DB 17; Length 169;
Best Local Similarity 28.8%; Pred. No. 0.67; Mismatches 60; Indels 24; Gaps 4;
Matches 40; Conservative 7;
QY 1 RLPSYSTSDSYPRGIGFKRPHTVAPW-----PAYPVTSYPLSQPDL 48
DB 27 RNPXRPRTTPXALPRP-----PSTPRPPXALPRPQAQPTXYPPL--QPRSPEDR 79
QY 49 LPIRRSPQLGSHRTSPSRDSCANSVASYENEEPCADEDEDYHNGYLVLDP 108
DB 80 PPPPLDSPCPRRHPPRPPRPPRMPA-----PPCVTPPPRQSGPHPPPTDAPSP 134
QY 109 STPATSTAPASAPLSTPG 127
DB 135 TPPKPLTPPHPTPALPTPG 153

RESULT 14
US-10-437-963-184670
; Sequence 184670, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/437.963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 184670
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(346)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_81641C.1.pep

US-10-437-963-184670
Query Match 10.3%; Score 110; DB 16; Length 346;
Best Local Similarity 24.0%; Pred. No. 1.9; Mismatches 41; Conservative 22; Indels 78; Gaps 5;
Matches 41; Conservative 22; Indels 78; Gaps 5;
QY 24 PHTVAPWPPAYP-----VTSYPLSQPDLPIPRSPQLGSHRTSPSRDSDGA 74
DB 163 PSSPSP-PPASPPSSSGAAAPTPPPASPPSTPTPPASPPASPPAPPS-----A 215
QY 75 NSVASYENEEPCADEDEDYHNGYLVLDPSTPATSTAPASAPALSTPGIRDSAFS 134
DB 216 NAPSQAQARNPSATSPRSR-----STAPAPSSSGAATPTAPAPRSSSSWS 262
QY 135 MESIDYVNVPESGESAASLDGSEYVNVSOELHPGAATPEPALSSQEA 185
DB 263 PTXTGSSASPPSSSSPSSGGGGGAGAPAGAPAPAT-PAGTNSRPA 312

RESULT 15
US-10-156-761-12405
; Sequence 12405, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: SATO, MASAHIRA
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156.761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12405
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12405

Query Match 10.3%; Score 110; DB 14; Length 584;
Best Local Similarity 25.3%; Pred. No. 3.6; Mismatches 50; Conservative 24; Indels 82; Gaps 11;
Matches 50; Conservative 24; Indels 82; Gaps 11;
QY 16 PRGIGFKRPHTVAPWPPAYPVTSYPLSQPDLPIPRSPQLGSHRTSPSRDSDGAN 75
DB 315 PTGAFGAPPTTATP-PPGPAP---WEPPAPPEPPAP-APAPAGRTPTDAGNIDSGS 369
QY 76 SVASYENEEPCADEDEDYHNGYLVLDPSTPATSTAPASAPALSTPGIRDSAFS 134
DB 370 G-----SCADDDLPYAPAPW-----KTPAPRTVPAVEVEVEKPPVRE---- 409
QY 135 MESIDYVNVPESGESAASLDG--SREYVNVSOELHPGAATPEPALSSQEAEBVEEG 192
DB 410 FDSVAEAVLAPESGSGDGAAGPLAPMARINEAVKQG--RIEAAAMAEEA---VHG 463
QY 193 A-----PDYENLOEL 202
DB 464 AATLGPHEPDPVLRREL 481

Search completed: November 21, 2004, 14:32:23
Job time : 143 secs


```
Query Match          11.3%; Score 121.5; DB 2; Length 253;
Best Local Similarity 23.6%; Pred. No.0.18;
Matches      61; Conservative    32; Indels   103; Gaps   13;
```

A/Accession: S26058
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-2090 <VON>

A:Cross-references: UNIPROT:P35658; EMBL:X64228; NID:G29652; PIDN:CAA4553.1; PID:G29653

Query Match 10.8%; Score 116; DB 2; Length 2000;

Best Local Similarity 24.9%; Pred. No. 5.3;

Matches 52; Conservative 20; Mismatches 85; Indels 52; Gaps 7;

QY 5 SYDSTSSDLSYPRGIGFQKRPHTVAPWPPVPPVTSYPLSQDLPPIRSPQ-PLGSGHR 63

DB 1440 SFGSGQNTSTV-----PSAPPTTATPLTSPFLTSPGSLSSATTPSLPMGAGRS 1492

QY 64 T-----PSSRRSDGANSVASYENEPACEDDEDEDYHNPGLVLPDSTPATST 115

DB 1493 TEATSSALPEKPGSEVASAASALLEEQOQSAQ-----LPQAPQTS 1535

QY 116 AAPSAALSTPGIRDS-----APSMESIDDVNVPESGSAEASLDGSRVNVNS 165

DB 1536 SYRKEBPVLAQPAVNSNGTASSTSLVALSAEATPATGVDA--RTEAVPPASSFSV--- 1590

QY 166 QELHPGAATKTEBPALSSQAEVEEBCAP 194

DB 1591 ----FGQTAVTAAAISSAGPAVETISSTP 1615

RESULT 6

hypoetical protein 45 - equine herpesvirus 2

C:Species: equine herpesvirus 2

C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #ext_change 09-Jul-2004

C:Accession: S55640

R:Jefford, B.A.R.; Watson, M.S.; Aird, H.C.; Perry, J.; Davison, A.J.

J. Mol. Biol. 249, 520-528, 1995

A:Title: The DNA sequence of equine herpesvirus 2.

A:Reference number: S55594; MUID:95302501; PMID:7783207

A:Accession: S55640

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-321 <TEL>

A:Cross-references: UNIPROT:Q66648; GB:U20824; NID:G695172; PIDN:AAC1383.1; PID:G695218

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1995

Query Match 10.6%; Score 113.5; DB 2; Length 321;

Best Local Similarity 25.0%; Pred. No. 0.83;

Matches 47; Conservative 23; Mismatches 61; Indels 57; Gaps 8;

QY 31 PPAVPPVTSYPLSQDLPPIRSPQPLGSHRTSSRRSDGANSVASYENB----- 84

DB 78 PPSHVPVPSKPK-----PRT-----RAGCAADSDSESEDEGDEDEETLMSQ 120

QY 85 ---PACEDADEDEDYHNPGLVLPDSTPATSTAAPALSTPGIRDSAFMSIDY 141

DB 121 DTPGSGSSSDSDDD-----QKLPATGCIKMPGYMS-RISDSSSSSSSSSD- 167

QY 142 VNVPSGSAEASLDGSR-----YVNVSGELHPGAATKTEPA-----ALSSQAE 186

DB 166 ---SSSSSSSDSESDGSTPEPDLIRQVTSGLAGVSPRAKPPAKKEVIVILSS 224

QY 187 EYEEBCAP 194

DB 225 ESDSEGE 232

RESULT 7

hypoetical protein XK84.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #ext_change 09-Jul-2004

C:Accession: T29018

R:Kirsten, J.

submitted to the EMBL Data Library, April 1995

A:Description: The sequence of C. elegans cosmid XK84.

A:Reference number: Z20553

A:Accession: T29018

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-801 <KIR>

A:Cross-references: UNIPROT:Q23635; EMBL:U23181; PIDN:AAC48204.1; GSPDB:GN00020; CESP:Z

A:Experimental source: strain Bristol N2; clone XK84

C:Gene: CESP:ZK84.1

A:Map position: 2

A:Introns: 22/2; 45/3; 108/1

C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology

Query Match 10.6%; Score 113.5; DB 2; Length 801;

Best Local Similarity 24.6%; Pred. No. 2.5;

Matches 58; Conservative 29; Mismatches 78; Indels 71; Gaps 11;

QY 4 GSYSSTSSDLSYPRGIGFQKRPHTVAPWPPA-----YPPVTSYPLSQ-----PPLL- 50

DB 157 GSGDAPASNSVTETLTATPAIILP-PPAEPRKIIIPASGVTSRQEGAGAPAPAV 215

QY 51 IPRSPQPLGSHRTSSRRS-----DG-----ANSVASYENEPACEDDEDED 95

DB 216 VEETPAETPAETPAETPAETASNANGTEPGYVDGTANAASAVAPAVETPAAPAE-- 273

QY 96 DYHNPGLVLPDSTPATSTAAPALSS-----TPGIRDSAPSMESIDVYVNPESGES 150

DB 274 -----TPAPATSSSEAPAPAPAEETPAETPAETPAETPAETPAETPAETPAET 317

QY 151 ABAASLDGSRVNVNVOELHPGAATKTEPA-----LSQAEVEEBCAP 195

DB 318 ATAPAP-----SSEADAAAPTDSASADTAAVDTSTALVDTSSSEHAETEAATD 365

RESULT 8

G-protein signaling regulator RGS3 - human

C:Species: Homo sapiens (man)

C:Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #ext_change 09-Jul-2004

C:Accession: S78089; S68436

R:Druey, K.

submitted to the EMBL Data Library, May 1995

A:Reference number: S78089

A:Accession: S78089

A:Molecule type: mRNA

A:Residues: 1-519 <DRU>

A:Cross-references: UNIPROT:Q49796; EMBL:U27655; NID:G1216368; PID:G1216369

R:Druey, K.M.; Blumer, K.J.; Kang, V.H.; Kehrl, J.H.

Nature 379, 742-746, 1996

A:Title: Inhibition of G-protein-mediated MAP kinase activation by a new mammalian gene

A:Reference number: A58012; MUID:96178495; PMID:8602223

A:Accession: S68436

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-355, 'K', 357-519 <DRW>

A:Cross-references: EMBL:U27655

A:Experimental source: tonsil

A>Note: the sequence from Fig. 1b is inconsistent with that from Fig. 1a in having 356-7

Query Match 10.5%; Score 112; DB 2; Length 519;

Best Local Similarity 24.3%; Pred. No. 1.9;

Matches 51; Conservative 22; Mismatches 79; Indels 58; Gaps 11;

QY 3 PGYSSTSSDLSYPRGIGFQKRPHTVAPWPPVPPVTSYPLSQ-----DLP----IPRP 55

DB 57 PSQEPPLSSKDSATSEB-----SPGPDAPPSKDVPPCGPPPADLSPQDLPACQ 107

QY 56 QPLGSHRTSSRRSDGANSVASYENEPACEDDEDED-----DYHNPGLVLP 106

DB 108 EPL--PHQDPLTLKLPALQ--SPTRDLPQGLDPPGQVSLPAKALREDTMSSGDLAA 163

QY 107 PDSTPATSTAAPALSTPGIR-DSAFMSIDVYVNPESGSAEASLDGSRVNVNS 165

DB 164 TGDPP-----AAP-RAFAVYFEVRLDSTYSQKA-----GAEQCGSGDEEDAEBA 206

QY 166 QELHPGAATKTEBPALSSQAEVEEBCAP 195

```

Db          207 BEVEEG-----EEGEDEDEDTSD 225
              :|: |          :| || | |:|

```

RESULT 9
 B40505
 hypothetical protein - suid herpesvirus 1 (strain Indiana-Funkhouser or Becker)
 C|Species: suid herpesvirus 1
 C|Date: 10-Apr-1992 #sequence_revision 10-Apr-1992 #text_change 09-Jul-2004
 C|Accession: B40505
 R|Cheung, A.K.
 J. Virol. 65, 5260-5271, 1991
 A|Title: Cloning of the latency gene and the early protein 0 gene of pseudorabies virus
 A|Reference number: A40505; MUID:91374576; PMID:1654441
 A|Accession: B40505
 A|Status: preliminary
 A|Molecule type: mRNA
 A|Residues: 1-1958 <CHE>
 A|Cross-references: UNIPROT:O69340; GB:M57505; NID:G334066; PIDN:AAA47468.1; P|D:G334066
 C|Superfamily: pseudorabies virus 1 nuclear antigen

Query Match	10.5%	Score	112	DB	2	Length	1958
Best Local Similarity	35.3%	Pred.	No.	9	2		
Matches	30	Conservative	5	Mismatches	26	Indels	24
						Gaps	3

QY 25 PMPPVIPPVTISYPPLSQPDLLPIPKRSQPGLGSS-----HRTPSRRDS 71
| | | | | | | | | | | | | | :
DB 483 PSPPRRP----PPLPPPPP PPPPPAGSARRRRRGCGGP PGGGRRRGCKRRA 538

```
Qy      72 DGANSVASYENEEFACEDADEDED 96
          :| : | | | | | :
Db    539 EGTEAAADADEE--EDGEDEDE 560
```

```

RESULT 10
T36729
Probable serine/threonine-specific protein kinase (EC 2.7.1.-) - Streptomyces coelicolor
C/Species: Streptomyces coelicolor
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 16-Aug-2004
C/Accession: T36729
R/Murphy, L.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1999
A/Reference number: Z21612
A/Accession: T36729
A/Status: translated from GB/EMBL/DDJ
A/Molecule type: DNA
A/Residues: 1-576 <MUR>
A/Cross-references: UNIPROT:O9XA04; EMBL:AL079308; PDB:CBM45227.1; GSPDB:GN00070; SCOP:
A/Experimental source: strain AJ(2)
C/Genetics:
A/Gene: SCOEDB:SCH69..30
C/Superfamily: protein kinase homology
C/Keywords: phosphotransferase; serine/threonine-specific protein kinase

Query Match      10 4%; Score 111.5; DB 2; Length 576;
Beet local Similarity 25.5%; Pred. No.23;
Matches 49; Conservative 12; Mismatches 74; Indels 57; Gaps 7;

```

Query Match	10.4%	Score	111.5	DB	2	Length	576
Best Local Similarity	25.5%	Pred.	No. 2.3				
Matches	49	Conservative	12	Mismatches	74	Indels	57
						Gaps	7

Qy	7	DSTSSDGLYPRGGLQFKPHNVAMWPAVYVPTVSYPRLSQDILLPHRSP-----QPLGGS	61
Db	300	DDVPTDPRRAG-----TPGVDSPPATPPT---PPAPRPWGTTPPAGSGGLDRSPGS	35
Qy	62	HRTSSKRDSDDGANSVSYNEEPPACEDADEDDDDYHNCGYLVLPLDSTPATSTLAPAP	12
Db	353	PGPPPTGPDSTPASP-----PPTPTVATITSPAP	38
Qy	122	ALSTPGIRDSAFSNESIDTVVNVPESGESAASLDGSRVYVNVSOELHPGAKTPEAALS	181
Qy	383	GL--PPASDDGWT--PSPPSGPTAPPAKSPASPSA-----PGTPRAPNGTH	427

QY	182	SQAEVEVEEGA	193
		:	
Db	428	SEEVPLAERPGA	439

RESULT 11

hypothetical protein al0962 - *Synechocystis* sp. (strain PCC 6803)
A:Species: *Synechocystis* sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #ext_change 09-Jul-2004
C:Accession: S74716
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yaeuda,
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
5.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S74716
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-477 <KAN>
A:Cross-references: UNIPROT:P72652; EMBL:D90901; GB:AB001339; NID:G1651897; PIDN:BAAL686
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match	10.4%	Score 11;	DB 2;	Length 477;
Best Local Similarity	24.9%	Pred. No. 2;		
Matches 48;	Conservative 24;	Mismatches 69;	Indels 32;	Gaps 57;

```
Qy      33  AYPPTSYTPPLSQDPLLPIPRSPQVLGGSHRTPSSRRSDSGANSVASYENEPACEDAE 92
      |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      256  AVPPV--QPRASQEPTEVQKVAIEBSSTAQTDDWDGDGEMDGLDPTSTEDSGPGELISDE 313
```

```

Qy      93 DEDDYHNP-----YLVLPDSTPATSTAAPSAPALSTPIGRDSPA 133
      : ||| : ||| : ||| :
Db      314 VSED--NPENDQKVIAVETVQIEQITLIEVPSENITAEAPAPFLIEEQIAQE-- 369

```

```
Qy      134 SMESIDDYVNVPESESAEASLDGSRKYVNTSQELHPGAAKTEPALSSQEAF-----EV 188
           | : | | | | : | : | : | : | : |
Db      370 --ETVVDEVIAPMGSTTAAIVETMPBELVPBETNINQGPDSSGVDEYERETETEIGEI 422
```

```
QY      189  EEEGAPDYENLQE  201
          |   |   |   |   |
Db      428  TEAIASDLEEVPE  440
```

RESULT 12
 T20884
 hypothetical protein FlAD7.2 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T20884
 R:Berk's, M.
 submitted to the EMBL Data Library, July 1996
 A:Reference number: Z19340
 A:Accession: T20884
 A:status: preliminary; translated from GB/EMBL/DDBT
 A:Molecule type: DNA
 A:Residues: 1-457 <MIL>
 A:Cross-references: UNIPROT:Q19452; EMBL:Z77658; PIDN:CA01156.1; GSPDB:GNO0023; CESP:FL
 A:Experimental source: clone FlAD7
 C:Genetics:
 A:Gene: CESP:FlAD7.2
 A:Map position: 5
 A:Introns: 31/2; 257/1; 287/3; 320/3; 343/3; 380/3

Query Match	10.3%	Score 11.0	DB 2	Length 457
Best Local Similarity	20.0%	Pred. No. 2.2		
Matches 57; Conservative	32;	Mismatches	80;	Indels 116; Gaps 10

QY 1 RLP-GSDVDSSTSDSLVPRGIQF-----KRPHTVAQPW-----PAPYEV 37
 28 RLPGSSGNGGTGSSFFRGGMMSNESSSAHDDKVIKKQETSSSSPEHRSPPRSKAPPS 87
 QY 38 TSYVPLSLQPLL-----PIRSP 55

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 14:10:14 ; Search time 194 Seconds

(without alignments)
602.067 Million cell updates/sec

Title: US-09-597-920b-4_COPY_31_233

Perfect score: 1071

Sequence: 1 RLPGSYDSTSSDLYPRGQI.....EAEVEBEGAPDYENQLHLN 203

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1046.5	97.7	262	1	LAT_HUMAN
2	683	63.8	242	1	LAT_MOUSE
3	683	63.8	242	2	CAD27390
4	670.5	62.6	241	1	Q80UR6
5	665	62.1	242	2	Q8C2E5
6	665	62.1	242	2	Q8C2E5
7	132	12.3	710	2	Q9CAL8
8	128	12.0	1296	2	Q6BR88
9	126.5	11.8	754	2	Q6PEI3
10	126.5	11.8	754	2	AAH58052
11	123.5	11.5	311	2	Q6L4T3
12	123.5	11.5	311	2	AAAT38083
13	122	11.3	704	1	NP14_PAT
14	121.5	11.3	253	2	Q9UFW0
15	121.5	11.3	367	2	Q9NEK0
16	121	11.3	619	1	Q96A17
17	121	11.3	699	1	NP14_HUMAN
18	121	11.1	700	2	Q9HUV3
19	119	11.1	703	2	Q6ZOK6
20	119	11.1	703	2	BAC97850
21	118.5	11.1	700	2	Q8CE21
22	118.5	11.1	1027	2	Q7Q2X2
23	118	11.0	686	2	Q874W2
24	116	10.8	1519	2	Q7Z3C4
25	116	10.8	2090	1	NP14_HUMAN
26	116	10.8	2091	2	Q75R47
27	116	10.8	2091	2	BAD07398
28	116	10.8	2093	2	Q86XD3
29	115	10.7	490	2	Q96BP7
30	115	10.7	490	2	Q96BS0
31	115	10.7	490	2	Q96BS0

32	115	10.7	1146	2	Q86XA7	Q86XA7 homo sapien
33	114.5	10.7	838	2	Q9VOA9	Q9VOA9 drosophila
34	114.5	10.7	917	2	Q81UQ1	Q81UQ1 homo sapien
35	114.5	10.7	1203	2	Q9ACK5	Q9ACK5 streptomyce
36	114	10.6	490	2	Q96BP6	Q96BP6 plasmodium
37	114	10.6	2183	2	Q6CEV2	Q6CEV2 yarrowia 11
38	113.5	10.6	321	2	Q66648	Q66648 equid herpe
39	113.5	10.6	801	2	Q23635	Q23635 caenorhabdi
40	113.5	10.6	1101	2	Q6BK14	Q6BK14 debaryomyce
41	113.5	10.6	1374	2	Q9VSU0	Q9VSU0 drosophila
42	113.5	10.6	1449	2	Q9U112	Q9U112 drosophila
43	113.5	10.6	1450	2	Q8IQB8	Q8IQB8 drosophila
44	113.5	10.6	2409	2	Q96OG6	Q96OG6 drosophila
45	113.5	10.6	2786	2	Q9VSU2	Q9VSU2 drosophila

ALIGNMENTS

RESULT 1	LAT_HUMAN	STANDARD	PRT	262 AA.
AC	Q43561; Q43919;			
AD	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	05-UTL-2004 (Rel. 44, Last annotation update)			
DE	Linker for activation of T cells (36 kDa phospho-tyrosine adaptor protein) (pp36) (p36-38).			
GN	Name=LAT;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	(1)			
RP	SEQUENCE FROM N.A. (ISOFORM LONG AND SHORT), AND SEQUENCE OF 32-47			
RP	AND 219-233.			
RC	TISSUE=Leukemia;			
RX	MEDLINE=98149346; PubMed=9489702;			
RT	Zhang W., Sloan-Lancaster J., Kitchen J., Tribble R.P., Samelson L.E.;			
RT	"LAT: the ZAP-70 tyrosine kinase substrate that links T cell receptor			
RT	to cellular activation."			
RL	Cell 92:83-92(1998).			
RN	(2)			
RP	SEQUENCE FROM N.A. (ISOFORM SHORT).			
RP	TISSUE=Thymus;			
RX	MEDLINE=98197173; PubMed=9529333;			
RT	Weber J.R., Oretavik S., Torgersen K.M., Danbolt N.C., Berg S.F.;			
RT	Ryan J.C., Tasken K., Imboden J.B., Vaage J.T.;			
RT	"Molecular cloning of the cDNA encoding pp36, a tyrosine-			
RT	phosphorylated adaptor protein selectively expressed by T cells and			
RT	natural killer cells."			
RL	J. Exp. Med. 187:1157-1161(1998).			
RN	(3)			
RP	SEQUENCE FROM N.A. (ISOFORM SHORT).			
RP	TISSUE=Colon;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RT	Strausberg R.L., Feigold E.A., Grouse L.H., Derge J.G.;			
RT	Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schaller G.D.;			
RT	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.;			
RT	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.;			
RT	Diatchenko L., Martusina K., Farmer A.A., Rubin G.M., Hong L.;			
RT	Stapleton M., Soares M.B., Bonaldi M.F., Caesvant T.L., Schetz T.E.;			
RT	Brownstein M.U., Usdin T.B., Toshiyuki S., Carninci P., Prange C.;			
RT	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.;			
RT	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.;			
RT	Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.;			
RT	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.;			
RT	Faney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.;			
RT	Whiting M., Madan A., Young A.C., Shaden A., Rodriguez S., Sanchez A.;			
RT	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.;			
RT	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.;			
RT	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.;			
RT	Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;			

RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Involved in T-cell antigen receptor (TCR) signal
 CC transduction pathway. May play an important role downstream of the
 CC activation of protein tyrosine kinases (PTKs).
 CC -1- SUBUNIT: When phosphorylated, interacts directly with the p85
 CC subunit of phosphoinositide 3-kinase and the SH2 domain of GRB2,
 CC GRAP, and PLC-gamma-1. Interacts indirectly with CBL, SOS, VAV,
 CC and SLP-76.
 CC -1- SUBCELLULAR LOCATION: Type III membrane protein (Potential).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=long;
 CC IsoId=043561-1; Sequence=Displayed;
 CC Name=short;
 CC IsoId=043561-2; Sequence=VSP_004303;
 CC -1- TISSUE SPECIFICITY: Expressed in thymus, T cells, nk cells, mast
 CC cells and, at lower levels, in spleen.
 CC -1- PTM: Phosphorylated on tyrosines by ZAP-70 and SYK.
 CC Phosphorylation leads to the recruitment of multiple signaling
 CC molecules. Is one of the most prominently tyrosine-phosphorylated
 CC proteins detected following TCR engagement.
 CC -1- MISCELLANEOUS: Engagement of killer inhibitory receptors (KIR)
 CC disrupts the interaction of PLC-gamma-1 with LAT and blocks target
 CC cell-induced activation of PLC, maybe by inducing the
 CC dephosphorylation of LAT.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF036906; AAC39637.1; -
 CC EMBL: AF036905; AAC39636.1; -
 CC EMBL: AJ223280; CA11218.1; -
 CC EMBL: BC011563; AAH11563.1; -
 CC Genew: HGNC:18974; LAT.
 CC DR MIM: 602354; -
 CC DR GO: GO:0005070; F:SH3/SH2 adaptor protein activity; TAS.
 CC DR GO: GO:0006955; P:Immune response; TAS.
 CC DR InterPro: IPR006359; LAT.
 CC DR PRINTS: PR01781; LATPROTEIN.
 CC KW Alternative splicing; Direct protein sequencing; Phosphorylation;
 CC Signal-anchor; Transmembrane.
 CC FT DOMAIN 1 4
 CC FT TRANSMEM 5 27
 CC FT 28 262
 CC FT DOMAIN 161 164
 CC FT DOMAIN 200 203
 CC FT DOMAIN 220 223
 CC FT MOD_RES 110 110
 CC FT MOD_RES 156 156
 CC FT MOD_RES 200 200
 CC FT MOD_RES 220 220
 CC FT MOD_RES 255 255
 CC FT VARSPIC 114 142
 CC FT SEQUENCE 262 AA; 27929 MW; BCD80AE7DCA64153 CRC64;
 CC Query Match 97.7%; Score 1046.5; DB 1; Length 262;
 CC Best Local Similarity 87.5%; Pred. No. 4.9e-58;
 CC Matches 203; Conservative 0; Mismatches 0; Indels 29; Gaps 1;

QY 1 RLPGSYDSTSDSLYPKGIQFKRPHVAPWPPVPTVSYPPLSQPDLRIPRSPQLGG 60
 Db |||||
 31 RLPGSYDSTSDSLYPKGIQFKRPHVAPWPPVPTVSYPPLSQPDLRIPRSPQLGG 90
 QY 61 SHRTSSRRSDGANSYASIEENE-----EACGDAD 91
 Db |||||
 91 SHRTSSRRSDGANSYASIEENEASGIRGAQGWGWSWTRLTPVSLPEEPACDAD 150
 QY 92 EDEDYHNPGLVYLPPSTATSTPAASAPALSTPGIRDSAFSMESIDYVYNVPESEGA 151
 Db |||||
 151 EDEDYHNPGLVYLPPSTATSTPAASAPALSTPGIRDSAFSMESIDYVYNVPESEGA 210
 QY 152 EASLDGSEYVYNVQELHPGAKTPEPALSSQAEVEEYGADPYENLQELN 203
 Db |||||
 211 EASLDGSEYVYNVQELHPGAKTPEPALSSQAEVEEYGADPYENLQELN 262
 RESULT 2
 LAT_MOUSE STANDARD; PRT; 242 AA.
 ID LAT_MOUSE 054557;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUN-2004 (Rel. 44, Last annotation update)
 DE Linker for activation of T cells (36 kDa phospho-tyrosine adaptor
 DE protein) (p36) (p36-38).
 GN Name=Lat;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thymus;
 RX MEDLINE=36149346; PubMed=9489702;
 RA Zhang W., Sloan-Lancaster J., Kitchen J., Trible R.P., Samelson L.E.;
 RT "LAT: the ZAP-70 tyrosine kinase substrate that links T cell receptor
 RL to cellular activation.";
 RL Cell 92:83-92(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Salivary gland;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
 RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshlyuk S., Carninci P., Prange C.,
 RA Raha S.S., Loguélano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.U., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguezes S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Kiryushinski M.I., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.U.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 CC and mouse cDNA sequences".
 CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Involved in T-cell antigen receptor (TCR) signal
 CC transduction pathway. May play an important role downstream of the
 CC activation of protein tyrosine kinases (PTKs) (By similarity).
 CC -1- SUBUNIT: When phosphorylated, interacts directly with the p85
 CC subunit of phosphoinositide 3-kinase and the SH2 domain of GRB2,
 CC GRAP, and PLC-gamma-1. Associates indirectly with CBL, SOS, VAV,
 CC and SLP-76 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type III membrane protein (Potential).
 CC -1- TISSUE SPECIFICITY: Thymus.
 CC -1- PTM: Phosphorylated on tyrosines by ZAP-70 and SYK.

Phosphorylation leads to the recruitment of multiple signaling molecules. Is one of the most prominently tyrosine-phosphorylated proteins detected following TCR engagement (by similarity).

- MISCELLANEOUS: Engagement of Killer Inhibitory Receptors (KIR) disrupts the interaction of PLC-gamma-1 with LAT and blocks target cell-induced activation of PLC, maybe by inducing the dephosphorylation of LAT (by similarity).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AF036907; AAC04054.1; -
 DR MGD; BC013337; AAH13337.1; -
 DR InterPro; IPR008359; LAT.
 DR PRINTS; PR01701; LATPROTEIN.
 KM Phosphorylation; Signal-anchor; Transmembrane.
 FT DOMAIN 1 4 Extracellular (Potential).
 FT TRANSMEM 5 28 Signal-anchor for type III membrane protein (Potential).
 FT DOMAIN 29 242 Cytoplasmic (Potential).
 FT DOMAIN 136 139 PLC-GAMMA-1-BINDING MOTIF.
 FT DOMAIN 175 178 GRB2-BINDING MOTIF 1.
 FT DOMAIN 195 198 GRB2-BINDING MOTIF 2.
 FT MOD_RES 113 113 Phosphotyrosine (by Tyr-kinases) (Potential).
 FT MOD_RES 132 132 Phosphotyrosine (by Tyr-kinases) (Potential).
 FT MOD_RES 175 175 Phosphotyrosine (by Tyr-kinases) (Potential).
 FT MOD_RES 195 195 Phosphotyrosine (by Tyr-kinases) (Potential).
 FT MOD_RES 235 235 Phosphotyrosine (by Tyr-kinases) (Potential).
 SQ SEQUENCE 242 AA; 26014 MW; 6AC25F7ABE61A5C1 CRC64;

Query Match 63.8%; Score 683; DB 1; Length 242;
 Best Local Similarity 68.9%; Pred. No. 2.6e-35;
 Matches 146; Conservative 16; Mismatches 38; Indels 12; Gaps 6;

QY 2 LFGSYDSTSSDGLYRGIGQFKRPHVAPMPA--YPPVTSYPPLSQPDLLIPRSPQPLG 59
 DB 33 LPVSYDSTSTESLYPRSLIKPPQITVPRTPAVSYPLVTSFPPLQPDLLIPRSPQPLG 92
 QY 60 GSHRTPSSRRDSDGANSVASYENEPACE--DADEDEDYHNPGLVLPDSTPATSTAA 117
 DB 93 GSHRPPSSQQNDSDANSVASYENQEPACKNVADDEDEDYR-GLVLPDPSSPAVAVPV 151
 QY 118 PSAPALSTPGIRDSAPSMESIDYVNVPSGSGAASLDGSEYVNVSOELHPGAKTTP 177
 DB 152 SSAPVPSNPNDLSDSAFVSCECDYVNVPSSESASLSLDGSEYVNVSPQOP-VTRAE 210
 QY 178 AALSQGEAE-EYEEG-----APDYENIQELN 203
 DB 211 ASVNSQVEBDEGEEGCVGDEAPDYENIQELN 242

RESULT 3
 CAD27390 PRELIMINARY; PRT; 242 AA.
 AC CAD27390;
 DT 02-MAR-2004 (TREMBlrel. 27, Created)
 DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
 DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
 DE Linker protein.
 GN LAT.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI TaxId=10090;
 RX [1] -
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/Ola;
 RA Aguado B., Nunez-Cruz S., Miazek A., Richeine S., Mura A.M.,
 RA Richeine M., Sainy D., He H.T., Malsen B., Malsen M.;
 RT "Spontaneous and exaggerated T helper type 2 immunity in mice with a
 RT point mutation in LAT.";
 RL Science 296:2036-2040(2002).
 DR EMBL; AJ438435; CAD27390.1; -
 FT CHAIN 1 242 POTENTIAL.
 SQ SEQUENCE 242 AA; 26014 MW; 6AC25F7ABE61A5C1 CRC64;

Query Match 63.8%; Score 683; DB 2; Length 242;
 Best Local Similarity 68.9%; Pred. No. 2.6e-35;
 Matches 146; Conservative 16; Mismatches 38; Indels 12; Gaps 6;

QY 2 LFGSYDSTSSDGLYRGIGQFKRPHVAPMPA--YPPVTSYPPLSQPDLLIPRSPQPLG 59
 DB 33 LPVSYDSTSTESLYPRSLIKPPQITVPRTPAVSYPLVTSFPPLQPDLLIPRSPQPLG 92
 QY 60 GSHRTPSSRRDSDGANSVASYENEPACE--DADEDEDYHNPGLVLPDSTPATSTAA 117
 DB 93 GSHRPPSSQQNDSDANSVASYENQEPACKNVADDEDEDYR-GLVLPDPSSPAVAVPV 151
 QY 118 PSAPALSTPGIRDSAPSMESIDYVNVPSGSGAASLDGSEYVNVSOELHPGAKTTP 177
 DB 152 SSAPVPSNPNDLSDSAFVSCECDYVNVPSSESASLSLDGSEYVNVSPQOP-VTRAE 210
 QY 178 AALSQGEAE-EYEEG-----APDYENIQELN 203
 DB 211 ASVNSQVEBDEGEEGCVGDEAPDYENIQELN 242

RESULT 4
 Q080UR6 PRELIMINARY; PRT; 242 AA.
 AC Q080UR6;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Linker for activation of T cells.
 GN Name=Lat;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI TaxId=10090;
 RX [1] -
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6NCR; TISSUE=Hematopoietic Stem Cell;
 RC MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
 RA Brownstein M.J., Ueidi T.B., Toshiyuki S., Carinici P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.D.,
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.D., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalios D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6NCR; TISSUE=Hematopoietic Stem Cell;
 RA Strauberg R.;
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC052340; AAH52340.1; -
 DR MGD; MGI:1342293; Lat.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro: IPR008359; LAT.
 DR PRINTS: PR01781; LATPROTEIN.
 DR SEQUENCE 242 AA; 25984 MW; 96955D5ACFA1F61 CRC64;
 SQ
 Query Match 63.3%; Score 678; DB 2; Length 242;
 Best Local Similarity 68.4%; Pred. No. 5.4e-35;
 Matches 145; Conservative 16; Mismatches 39; Indels 12; Gaps 6;
 QY 2 LPSGYDSTSSDLSYPRGIQFKRPHVAWPPA--YPPVTSPPLSOPDLPIPPSPPLG 59
 DB 33 LPSVYDSASTSLYPRSLIKRPQITVTRPAVSPLVTSPPRQDLPIPPSPPLG 92
 QY 60 GSHRTPSRSDSDGANSVASYENEPACE--DADEDEDYHNPGLVLPDSTPATSTAA 117
 DB 93 GSHRMPSSQNSDDANSVASYENEPACKNVDADEDDYRN-GYLVLVLPSSPAAPV 151
 QY 118 PSAPALSTPGIRDSAFMESIDYVNPVPSGESASALDGSREYVNSQELHPGAKTER 177
 DB 152 SSAPVPSNPDLGDSAFVSESCEDYVNPVSESESAASLDGSREYVNSPEQAP-VTRAEL 210
 QY 178 AALSQGEAF-EVEBERG-----APDYENLOELN 203
 DB 211 ASVNSQVEDEBERGVDGEAPDYENLOELN 242

RESULT 5
 ID LAT_RAT STANDARD; PRT; 241 AA.
 AC 070601;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Linker for activation of T cells (36 kDa phospho-tyrosine adaptor protein) (p36) (p36-38).
 DE Name=Lat;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NC NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PVG; TISSUE=Natural killer cells;
 RX MEDLINE=98197173; PubMed=9529333;
 RA Weber J.R., Orestavik S., Torgersen K.M., Danbolt N.C., Berg S.F., Ryan J.C., Tasken K., Imboden J.B., Vaage J.T.;
 RT "Molecular cloning of the cDNA encoding p36, a tyrosine-phosphorylated adaptor protein selectively expressed by T cells and natural killer cells."
 RT phosphotyrosine kinase.
 RL J. Exp. Med. 187:1157-1161(1998).
 CC -1- FUNCTION: Involved in T-cell antigen receptor (TCR) signal transduction pathway. May play an important role downstream of the activation of protein tyrosine kinases (PTKs) (By similarity).
 CC -1- SUBUNIT: When phosphorylated, interacts directly with the p85 subunit of phosphoinositide 3-kinase and the SH2 domain of GRB2, GAP, and PLC-gamma-1. Associates indirectly with CBL, SOS, VAV, and SHP-2 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type III membrane protein (Potential).
 CC -1- TISSUE SPECIFICITY: NK cells.
 CC -1- PTM: Phosphorylated on tyrosines by ZAP-70 and SYK.
 CC Phosphorylation leads to the recruitment of multiple signaling molecules. Is one of the most prominently tyrosine-phosphorylated proteins detected following TCR engagement (By similarity).
 CC -1- MISCELLANEOUS: Engagement of killer inhibitory receptors (KIR) disrupts the interaction of PLC-gamma-1 with LAT and blocks target cell-induced activation of PLC, maybe by inducing the

CC dephosphorylation of LAT (By similarity).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; AJ001184; CAA04577.1; -
 DR RGD; 620802; Lat.
 DR InterPro: IPR008359; LAT.
 DR PRINTS: PR01781; LATPROTEIN.
 KW Phosphorylation; Signal-anchor; Transmembrane.
 FT DOMAIN 1 4
 FT TRANSMEM 5 28
 FT 29 241
 FT DOMAIN 217 224
 FT DOMAIN 136 139
 FT DOMAIN 175 178
 FT DOMAIN 195 198
 FT MOD_RES 113 113
 FT MOD_RES 132 132
 FT MOD_RES 175 175
 FT MOD_RES 195 195
 FT MOD_RES 234 234
 FT MOD_RES 234 234
 SQ SEQUENCE 241 AA; 26192 MW; FBD5C4DA1B8C4275 CRC64;
 Query Match 62.6%; Score 670.5; DB 1; Length 241;
 Best Local Similarity 67.3%; Pred. No. 1.6e-34;
 Matches 142; Conservative 16; Mismatches 42; Indels 11; Gaps 5;
 QY 2 LPSGYDSTSSDLSYPRGIQFKRPHVAWPPA--YPPVTSPPLSOPDLPIPPSPPLG 59
 DB 33 LPSVYDSASTSLYPRSLIKRPQITVTRPAVSPLVTSPPRQDLPIPPSPPLG 92
 QY 60 GSHRTPSRSDSDGANSVASYENEPACE--DADEDEDYHNPGLVLPDSTPATSTAA 117
 DB 93 GSHRMPSSQNSDDANSVASYENEPACKNVDADEDDYRN-GYLVLVLPSSPAAPV 151
 QY 118 PSAPALSTPGIRDSAFMESIDYVNPVPSGESASALDGSREYVNSQELHPGAKTER 177
 DB 152 SSAPVPSNPDLGDSAFVSESCEDYVNPVSESESAASLDGSREYVNSQAP-VTRAEL 210
 QY 178 AALSQGEAF-EVEBERG-----APDYENLOELN 203
 DB 211 ASVNSQVEDEBERGVDGEAPDYENLOELN 241

RESULT 6
 ID Q8C2E5 PRELIMINARY; PRT; 242 AA.
 AC Q8C2E5;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Mus musculus 2 days neonate thymic cells cDNA, RIKEN full-length enriched library, clone:R330025I08 product:linker for activation of T cells, full insert sequence.
 DE Name=Lat;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

```

RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=99379253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN PANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RA The PANTOM Consortium;
RT "The RIKEN Genome Exploration Research Group Phase I & II Team;
  Analysis of the mouse transcriptome based on functional annotation of
  60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=20493974; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
  Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtration of cap-trapper-selected cDNAs to
  prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1517-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
  Kono H., Akiyama J., Nishi K., Katsunai T., Taahiro H., Itoh M.,
  Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
  Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
  Fujitake S., Inoue K., Togawa Y., Izawa M., Ohara E., Wachihi M.,
  Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
  Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
  sequencing pipeline with 384 multiplexed capillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=20530913; PubMed=11076861;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
  Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
  Hayashida K., Hayatsu N., Hiramoto K., Hiraka T., Hirozane T.,
  Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kaubawa T.,
  Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
  Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
  Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
  Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
  Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
  Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RT Submitted (Apr-2002) to the EMBL/Genbank/DBJ databases.
RL EMBL; AK088763; BAC40554.1; -.
DR MGD; MGI:1342293; Lat.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR008359; LAT.
DR PRINTS; F01761; LATPROTEIN.
SQ SEQUENCE 242 AA; 26076 MW; D7A74C50CFE5AB35 CRC64;

```

```

QY 60 GSHRTSSRRSDGANSVASYENEPACE--DADDEDYINPGYLVLPDSTPATSTRA 117
DB 93 GSHRPSSQONSVDANSVASYENEPACKNDVDEVDYPR-GYLVLPDSSPAVPV 151
QY 118 PSAPALSTPGIRDSAFSMESIDYVNPESGESASASIDGSEYVNSOELHPGAKTER 177
DB 152 SSAPVPSPNDLDSAFVSESCEDYVNPESSESASASIDGSEYVNSPBPQP-VTRAE 210
QY 178 AALSQGEAE-EVEEG-----APDYENIQLIN 203
DB 211 ASVNSQVEDEBGECEGVDEEAPDYENIQLIN 242

RESULT 7
Q9CAL8 PRELIMINARY; PRT; 710 AA.
ID Q9CAL8;
AC Q9CAL8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein F24J13.3.
GN Name=F24J13.3;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
  Maiti R., Roming C.M., Coo H., Fujii C.Y., Uterback T.R.,
  Barnstead M.E., Bowman K.C., White O., Nierman W.C., Frazer C.M.;
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Town C.D., Kaul S.;
RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.
CC -1-SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AC010796; AAG52479.1; -.
DR PIR; D96728; D96728.
DR HSP; P36897; IIAS.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot. Kinase.
DR InterPro; IPR008271; Ser_Thr_Pkin_AS.
DR Pfam; PF00069; PKinase; 1.
DR PROSITE; PS00001; Prot. Kinase; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; Hypothetical protein; kinase;
KW Serine/threonine-protein kinase; transferase.
SQ SEQUENCE 710 AA; 75371 MW; 8B4420A8D8F3A6 CRC64;

Query Match 12.3%; Score 132; DB 2; Length 710;
Best Local Similarity 31.6%; Pred. No. 3;
Matches 49; Conservative 6; Mismatches 52; Indels 48; Gaps 6;

```

RESULT 8

Q6BR88 PRELIMINARY; PRT; 1296 AA.

AC Q6BR88; 01-OCT-2004 (TREMBlrel. 28, Created)

DT 01-OCT-2004 (TREMBlrel. 28, Last sequence update)

DT 01-OCT-2004 (TREMBlrel. 28, Last annotation update)

DE Similar to CA5851|CAK12 Candida albicans CAK12.

GN OFNames=DEHA0D19833g;

OS Debaryomyces hansenii (Yeast) (Torulasporea hansenii).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetaceae; Saccharomycetaceae; Debaryomyces.

OX NCBI_TaxId=4959;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CB8767;

RG GENOLEVURES;

RA Lafontaine I., de Montigny J., Marck C., Neugebäude C., Talia E.,

RA Goffard N., Frangoul L., Aigle M., Anthouard V., Babour A., Barde V.,

RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,

RA Boismere A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,

RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Giropi A.,

RA Hantaye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,

RA Kerrest A., Kozul R., Lemaire M., Lesur I., Ma L., Muller H.,

RA Niclaud J.M., Nikolaki M., Oztas S., Ozier-Kalogeropoulos O.,

RA Pellenz S., Potter S., Richard G.F., Straub M.L., Suleau A.,

RA Svenne D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,

RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,

RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,

RA Wincker P., Soulier J.L.;

RT "Genome evolution in yeasts."

RL Nature 430:35-44 (2004).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CB8767;

RG GENOSCOPE;

RA Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.

RL EMBL; CR82136; CAG87456.1; -

SQ SEQUENCE 1296 AA; 14462 MW; CEAS4C1C112C2CCD CRC64;

Query Match 12.0%; Score 128; DB 2; Length 1296;

Best Local Similarity 24.1%; Pred. No. 11;

Matches 55; Conservative 25; Mismatches 80; Indels 68; Gaps 9;

QY 3 PGSYDSTSSDSLYRGIQFKRPHTVAPWPPAYPPV-----TSYPPLSQPDLLPIRSPQP 57

DB PSSKSHKNDPMVNRGVSMK---VTAKKSSSSKVDLHSTSNTPKQPQEV-----STOP 830

QY 58 LGGSHRPPSSRRDDGANSVASYENEPACEDADEDDYHNPGLVVLPPSTPATSTPA 117

DB TKTSHNRSSTSTNNAAVKGIVPEQ-----YLPPLPININTNTLV 873

QY 118 PSA-----PALSTGIRDSAF-----SMESIDYVNVPESGESASALDGSREYVNVSGEL 168

DB PPAGSGNSPLKVAAGABGRKHAFTARAKSVGHNKXSGRSNRA-----QKP 922

QY 169 HPGAAKTEPALSSQEAEEVEEGAP-----DYENKQEL 202

DB 923 HP-----TLPNALAGNSDEIDEGSTRDDGFDDVTLDDEVYQEIPLQ 966

RESULT 9

Q6PE13 PRELIMINARY; PRT; 754 AA.

AC Q6PE13; 05-JUL-2004 (TREMBlrel. 27, Created)

DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)

DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)

DE Hypothetical protein zgc:63484.

GN Name=zgc:63484;

OS Brachydanio rerio (Zebrafish) (Danio rerio).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;

OC Cyprinidae; Danio.

OX NCBI_TaxId=7935;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=AB; TISSUE=Whole body;

RX MEDLINE=2388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heile F.,

RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Udén T.B., Tsouhuki S., Carninci P., Prange C.,

RA Rana S.S., Loggiano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunnarone P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.U., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywinski M.I., Skalka U., Smaltus D.E., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=AB; TISSUE=Whole body;

RA Strausberg R.;

RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC058052; AAH58052.1; -

DR InterPro; IPR009043; RNA_pol_sigma.

DR InterPro; IPR004018; RPEL_repeat.

DR Pfam; PF02755; RPEL; 2.

DR SMART; SM00707; RPEL; 3.

KW Hypothetical protein.

SQ SEQUENCE 754 AA; 84183 MW; DSDDE29D4C1E7ADE CRC64;

Query Match 11.8%; Score 126.5; DB 2; Length 754;

Best Local Similarity 26.3%; Pred. No. 7.1;

Matches 50; Conservative 25; Mismatches 70; Indels 45; Gaps 10;

QY 16 PRGIQFKRPHTVAPWPPAYPPVTSYPLSQP-----DL-----PIPRSPQLGSGHRT 65

DB 422 PGGYQL-----PPPPSPPLPTHTIPSPRAHTNHLHQHSYYP-LPQPL-PVHFD 472

QY 66 SSRDSDGANSVASYENEPACEDADEDDYHNPGLVVLPPSTPATSTPAASAPALST 125

DB 473 PSPPEPPADDEDDYSDEE-----EEEDDEDEDEPP-----PHLPSPQQLPELPPSR 523

QY 126 PGIRDAFMSIDYVNVPESGESASALDGSREYVNVSGELHPGAKE-PAALSSQE 184

DB 524 CLVGEISVSV-----IPEGNNSSREED-----BEDQHPESSDSDGVLVYKDE 567

QY 185 AEEVEEGAP 194

DB 568 SDEDEEDDSP 577

RESULT 10

AAH58052 PRELIMINARY; PRT; 754 AA.

AC AAH58052; 24-MAY-2004 (TREMBlrel. 27, Created)

DT 24-MAY-2004 (TREMBlrel. 27, Last sequence update)

DT 24-MAY-2004 (TREMBlrel. 27, Last annotation update)

DE Hypothetical protein zgc:63484.

GN Brachydanio rerio (Zebrafish) (Danio rerio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OC NCBI_TaxId=7955;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=AB; TISSUE=whole body;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Struhsberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Scheinen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Locuelli N.A., Peters G.J., Abramson R.D., Mullighan S.J.,
 RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gamarine P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN (12)
 RP SEQUENCE FROM N.A.
 RC STRAIN=AB; TISSUE=whole body;
 RA Struhsberg R.;
 RL Submitted (SEP-2003) to the EMBL/Genbank/DBJ databases.
 DR EMBL; BC058052; AAHS052.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 754 AA; 84183 MW; D5DED29D4C1B7ADE CRC64;

Query Match 11.8%; Score 126.5; DB 2; Length 754;
 Best Local Similarity 26.3%; Pred. No. 7.1;
 Matches 50; Conservative 25; Mismatches 70; Indels 45; Gaps 10;

QY 16 PRGLOPKRPHVAVPMPAPVPTSYRPLSQP-----DLI-----DIPRSPQPLGSGSHRTP 65
 DB 422 PGGYL-----PPPPSPPLPTHTIPSPRPAHTHLLHQSYYR-LPQYL-FVNDP 472
 QY 66 SSRSDGANSVASYENEPACEDADEDDYHNGYLVLPDSTPATSTAPSPALST 125
 DB 473 PSPPEPPARDEDDYDSE--EEBEDDEDEPP-----PDHLPSQSQPLPRSR 523
 QY 126 PGIRSAASMSIDYVNVPSGESASASLDGSRVYVNSQELHGAATK-PALSSQE 184
 DB 524 CLVGLSLVSV-----IPEGNNSEEBED-----EEDQHPESDSDGFLVYKDE 567
 QY 185 AEEVEEGAP 194
 DB 568 SDEDEDDSP 577

RESULT 11
 O6L4T3 PRELIMINARY; PRT; 311 AA.
 AC O6L4T3;
 DT 05-JUN-2004 (TRENBLrel. 27, Created)
 DT 05-JUN-2004 (TRENBLrel. 27, Last sequence update)
 DE 05-JUN-2004 (TRENBLrel. 27, Last annotation update)
 DR Hypothetical protein P0478F09.11.
 GN Name=P0478F09.11;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Euphorbiaceae; Oryzae; Oryza.
 OC NCBI_TaxId=39947;
 RN (1)

RP SEQUENCE FROM N.A.
 RA Chow T.-Y., Hsiang Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
 RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
 RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
 RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kuo P.-I., Lee M.-C.,
 RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
 RA Wu H.-P., Shaw J.-F.;
 RL Submitted (MAY-2004) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AC130610; AAT38083.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 311 AA; 33620 MW; 56CAF49DA2B3886B CRC64;

Query Match 11.5%; Score 123.5; DB 2; Length 311;
 Best Local Similarity 23.1%; Pred. No. 3.9;
 Matches 46; Conservative 30; Mismatches 84; Indels 39; Gaps 6;

QY 30 WPPAPVPTSYRPL-----SQPLLPTRSPQPLGSHRPPSSRRSDGANSVASYENEP 86
 DB 53 WAPAPLTTTSWADVDDDDDDYFATTAAPRPRVGTHTHHAADHDDHDEQAALDELESE 112
 QY 87 CEDADEDDYHNGYLVLPDSTPA-----TSTAAPSPALSTPG----- 127
 DB 113 DEEVDDADDEH-----ETEDATPAEPANMKAAPAPPRDTERQSKKEKKELEL 168
 QY 128 ---IRDSAFMSIDYVNVPSGESASASLDGSRVYVNSQELHGAATKPEPALSSQE 184
 DB 169 DAIIELELSSKNNDACNETNGKGAQAADG-----NKEGAPADESK-----SSK 218

QY 185 AEEVEEGAPDENLELN 203
 DB 219 KKAKDKSAKAKETQELN 237

RESULT 12
 AAT38083 PRELIMINARY; PRT; 311 AA.
 AC AAT38083;
 DT 01-JUN-2004 (TRENBLrel. 27, Created)
 DT 01-JUN-2004 (TRENBLrel. 27, Last sequence update)
 DE 01-JUN-2004 (TRENBLrel. 27, Last annotation update)
 DR Hypothetical protein P0478F09.11.
 GN P0478F09.11.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Euphorbiaceae; Oryzae; Oryza; Oryza sativa.
 OC NCBI_TaxId=39947;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Chow T.-Y., Hsiang Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
 RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
 RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
 RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kuo P.-I., Lee M.-C.,
 RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
 RA Wu H.-P., Shaw J.-F.;
 RL Submitted (MAY-2004) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AC130610; AAT38083.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 311 AA; 33620 MW; 56CAF49DA2B3886B CRC64;

Query Match 11.5%; Score 123.5; DB 2; Length 311;
 Best Local Similarity 23.1%; Pred. No. 3.9;
 Matches 46; Conservative 30; Mismatches 84; Indels 39; Gaps 6;

QY 30 WPPAPVPTSYRPL-----SQPLLPTRSPQPLGSHRPPSSRRSDGANSVASYENEP 86
 DB 53 WAPAPLTTTSWADVDDDDDDYFATTAAPRPRVGTHTHHAADHDDHDEQAALDELESE 112
 QY 87 CEDADEDDYHNGYLVLPDSTPA-----TSTAAPSPALSTPG----- 127
 DB 113 DEEVDDADDEH-----ETEDATPAEPANMKAAPAPPRDTERQSKKEKKELEL 168

Oy	128----	IDSAPMSISIDYVNVPSGSEAEKSLDSSREYVNSOELHFGCAKTEPAALASOG	18
Dd	169	DATLEBLELSSSNDAQNETKGKAEGAAADGE-----NKGAFAPEAK-----SKK	218
Oy	185	AEEVEEGAPDYENIQELN 203	
Dd	219	KKAKDKSAKEAKEQTGLN 237	
<hr/>			
RESULT 13			
NP14_RAT			
ID NP14_RAT	STANDARD;	PRT;	704 AA.
AC P41777;			
DT 01-NOV-1995 (Rel. 32, Last sequence update)			
DT 01-NOV-1995 (Rel. 44, Last annotation update)			
DE Nucleolar phosphoprotein p130 (Nucleolar 130 kDa protein) (140 kDa nucleolar phosphoprotein) (Nopp140) (Nucleolar and coiled-body phosphoprotein 1).			
DE Name=NOL1;			
OS Rattus norvegicus (Rat).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX NCBI_Taxid=101116;			
RN [1]			
RP SEQUENCE FROM N.A., AND SEQUENCE OF 232-309 AND 563-601.			
RC TISSUE=Liver;			
RX MEDLINE=922323542; PubMed=1623516;			
RA Meier U.T.; Blobel G.;			
RT "Nopp140 shuttles on tracks between nucleolus and cytoplasm.";			
RL Cell 70:127-138(1992).			
RN [2]			
RP INTERACTION WITH NOPs AND FIBRILLARIN.			
RX MEDLINE=20143579; PubMed=10679015.			
RA Yang Y., Isaac C., Wang C., Dragoin P., Pogacic V., Meier U.T.;			
RT "Conserved composition of mammalian box H/ACA and box C/D small nucleolar ribonucleoprotein particles and their interaction with the common factor Nopp140."			
RL Mol. Biol. Cell 11:567-577(2000).			
CC -!- FUNCTION: Related to nucleogenesis, may play a role in the maintenance of the fundamental structure of the fibrillar center and dense fibrillar component in the nucleolus. It has intrinsic GTPase and ATPase activities. May play an important role in transcription catalyzed by RNA polymerase I (By similarity').			
CC -!- SUBUNIT: Interacts with DKC1/NAP57, NOP5/NAP65 and fibrillarin.			
CC -!- SUBCELLULAR LOCATION: Shuttles on curvilinear tracks between nucleolus and cytoplasm. These tracks extend from the dense fibrillar component of the nucleolus across the nucleoplasm to a limited number of nuclear pore complexes.			
CC -!- PTM: Undergoes rapid and massive phosphorylation/dephosphorylation cycles on CK2 and PKC sites. Nopp140 is one of the mostly phosphorylated proteins in the cell.			
CC -!- SIMILARITY: Contains 1 Lish domain.			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).			
CC -----			
DR EMBL, M94287; AAAA1718.1; -			
DR EMBL, M94288; AAAA1719.1; -			
DR PIR, B42680; B42680.			
DR InterPro, IPR006594; Lish.			
DR InterPro, IPR007718; SRP40_C.			
DR InterPro, IPR003993; Treacle.			
DR Pfam, PF05022; SRP40_C_1.			
DR PRINTS, PR01503; TREACLE.			
DR PROSITE, PS50896; LISH_1.			
KW ATP-binding; Direct protein sequencing; GTP-binding; Nuclear protein;			

Query Match	Best Local Similarity	Score 122, DB 1, Length 704;
Query	3 PGSYDS-----TSSDLYPRGIGQFKRPHTVAFM-----PP-----AYPV 37	11.3%; Score 121.5; DB 2; Length 253;
Db	12 PGSNESHHSRPEELPTPEPDPF-----PRLPLPWASHHGRGSPRPSLYSEAWPLT 64	23.6%; Pred. No. 4.1; Mismatches 63; Indels 103; Gaps 13;
Query	38 TSYPLPSQDILLPTPRSFQPLGGSHRTTPSSRRSDGANSVASYNEBPACDADDEDY 97	

```
Db      65 TSSAPGCEPALPCEHPEFPSCGS---PALREDLAAATPAS----- 103
QY      98 HNPGLVLP-DSTPAT---TAAP-----SAPALSTPGIRDSAFSME----- 136
Db      104 ---PLVLLPLETRPEBPQPSAPHPVPLSAAVPEGSPDRKQSRSLIALSGLEKL 159
QY      137 -----SIDDYVNVPESGESAB-----ASLDGSRREYVNVSOELHPGAaktePALSSQ 183
Db      160 KTVTSGSIQPVQAQAGQMDTKRLKDSAVLDQSAKYTHLTHD-----ELISLLIQ 211
QY      184 EAEVEEBCAPDYEINLOEL 202
Db      212 RERELSQRD---EHVQEL 226

RESULT 15
Q9NEKO PRELIMINARY; PRT; 367 AA.
ID Q9NEKO
AC Q9NEKO;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, last annotation update)
DE Hypothetical protein Y116F11B.8.
GN Name=Y116F11B.8;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
   investigating biology."
RL Science 282:2012-2018 (1998).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Sulston J.E., McElay K.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL132943; CAB81977.1; -.
DR WormPep; Y116F11B.8; CE24165.
KW Hypothetical protein.
SQ SEQUENCE 367 AA; 40617 MW; DBEB05068359F0FB CRC64;

Query Match      11.3%; Score 121.5; DB 2; Length 367;
Best Local Similarity 25.9%; Pred. No. 6.3;
Matches 43; Conservative 18; Mismatches 68; Indels 37; Gaps 4;

QY      36 PVTSTPPLSQPDLPLIPRSPOPLGSGHRTSSRRSDGANSVASYENEPACEDADEDED 95
Db      196 PVTSLPPPTDP-----PSTLGPMETPPNPGTSSAATTSAP----- 231
QY      96 DYHNGYLVLPDSTPATSTAPASAPALSTPGIRDSAFSMEISIDDYVNVPESGESABASL 155
Db      232 -----VTSFAPTEPASTLGPVVI PMNSPGTSSKATTSATVSLVLGPTTFSSAPQAL 283
QY      156 DGSREYVNVSOELHPGAaktePALSSQAEVEE---GAPYE 197
Db      284 DSGSLA-ADPSTDPSPSISSTDPPTYSSTTAEDYDEEDDKGDDEYD 328
```

Search completed: November 21, 2004, 14:20:03
Job time : 197 secs

THIS PAGE BLANK (USPTO)